VOL. 8 No. 11, p. 321-352 NOVEMBER 1946

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
N.V. PHILIPS' GLOEILAMPENFABRIEKEN

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, HOLLAND

X-RAY FLUOROSCOPY WITH ENLARGED IMAGE

by G. C. E. BURGER, B. COMBEE and J. H. van der TUUK.

621.386.8:
616-073.755:
620.179.152

The intrinsic blurring of a fluoroscope screen (for instance about 1 mm) interferes with the observation of small details showing little contrast. By placing the object at a distance from the screen and close to the focus an enlarged image is obtained in which the (constant) screen blurring has relatively less effect. In order to prevent the geometrical blurring, which increases rapidly upon enlargement, from spoiling the desired effect, X-ray tubes with a very small focus must be used for enlarged fluoroscopy. Upon enlargement all the contrasts in the image also increase considerably, due to the fact that with a large distance between object and screen practically no secondary, scattered radiation falls on the screen. By the improvement in definition and contrast certain groups of otherwise invisible details become perceptible, while at the same time the enlarged images can be studied more quickly and with less fatigue. In certain cases the enlargement is accompanied by a significant loss in (primary) screen brightness. The detrimental effect of this on the perceptibility of details is analysed. In medical diagnostics the enlargement makes it necessary to limit the time a patient is kept under observation. Phantom tests and a small series of examinations of lung patients have led to the conclusion that enlargement may well constitute an appreciable advance in diagnostics. The technique to be followed in this case has to be further worked out. In a separate section a description is given of the Philips X-ray tubes which were specially developed for the experiments in question. They work with electrostatic focusing and provide a choice of two foci, one normal focus of 2.0 or 1.2 mm width and a fine focus of 0.3 mm width.

Blurring of the image in fluoroscopy

The shadow picture obtained in X-ray screening of any object is always affected by a certain blurring, which may be ascribed two to causes. In the first place the focus of the X-ray from which the projecting rays are emitted is not a point. The finite width of the focus causes a "geometrical blurring" O_g (half shadow width), which according to *fig. 1* is given by

$$O_g = \frac{b}{a} f, \quad \ldots\ldots\ldots \quad (1)$$

where a and b are respectively the distance between object and focus and that between object and screen. In the second place the mechanism of the excitation of the light on the fluorescent screen causes a blurring of the image. Every fluoroscope screen has for this reason an intrinsic blurring, the "screen blurring", O_s, which for a normally good screen may amount, for example, to 0.6 to 1.0 mm. The geometrical blurring is usually smaller. For instance in the case of normal lung screening $a \approx 50$ cm, $b \approx 20$ cm, and foci of for instance 1.2 mm width are used; for such a case $O_g \approx 0.5$ mm.

In the testing of materials as well as in medical diagnostics the object in fluoroscopy is to be able to observe very small objects, often with little contrast (minute holes or cracks, errors in assembly of the piece of work; disease nuclei in the lungs, small ulcers in the stomach, etc.) The blurring of the image will be a handicap in such cases. We shall explain this further, but for the present it may be stated roughly that the perceptibility of

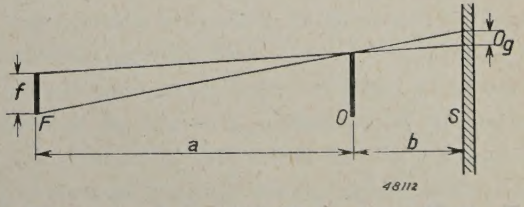


Fig. 1. Due to the finite width of the focus (F) the X-ray shadow picture of the object (O) on the fluoroscope screen (S) has a geometrical blurring O_g.

a detail with little contrast will become doubtful as soon as the detail is no larger than the blurring of the image. With the above mentioned numerical values of the blurring, therefore, all kinds of details (sometimes important ones) smaller than about 1 mm would escape observation.

Principle of enlarged fluoroscopy

It is obvious that only in case of necessity will such a limitation be tolerated. There now exists a method, which in principle is very simple, to improve the perceptibility of small details, namely fluoroscopy with enlarged image. For that purpose the arrangement of fig. 1 is modified in such a way that b becomes several times as large as a; see *fig. 2*. It may be read off from the figure that the object is projected on the screen enlarged by a factor

$$v = \frac{b + a}{a} \quad \ldots\ldots\ldots \quad (2)$$

Since all details are enlarged in the same proportion (or, with a thick object, in approximately the same proportion), while the intrinsic screen blurring remains unaltered, small details then have

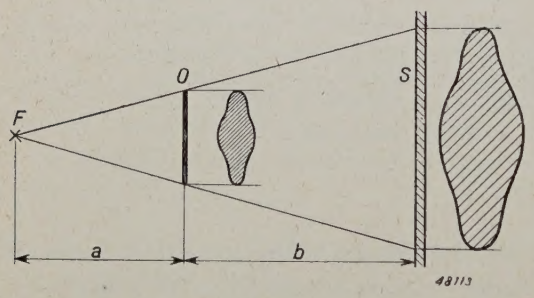


Fig. 2. By making the distance b between object (O) and screen (S) large enough, compared with the distance a between object and focus (F), a more or less enlarged fluoroscopy image of the object can be obtained.

a greater chance to show up above the screen blurring. This would indeed mean that with the change in the ratio of distance b/a mentioned the geometrical blurring would increase considerably (equation (1)) and therefore neutralize the advantage gained. It is possible, however, to compensate the increase of the factor b/a in equation (1) by making the width of the focus f correspondingly smaller. For the practical application of enlarged fluoroscopy therefore special X-ray tubes with a very small focus are desired.

In the following we shall replace this rough picture with a more quantitative formulation and also discuss the influence of various other factors involved in enlarged screening. It will be found that the situation for the testing of materials is somewhat different from that for medical diagnostics. The possibilities of medical applications, which have been studied carefully by us in recent years, will be dealt with in somewhat more detail. At the same time the X-ray tubes especially developed by Philips in connection with the magnification technique will be described.

The improvement attainable by magnification

The critical size of detail

It is unnecessary to prove that a larger and sharper image is more pleasant and offers more certainty of judgement. But it may be asked whether in the enlarged, relatively sharper image details can indeed be observed which were formerly invisible. In order to answer this question it is first necessary to consider the detrimental effect of the blurring.

To begin with, we condiser the case where b is very small; no enlargement is thus employed ($v \approx 1$) and we also assume a point focus, so that the geometrical blurring is zero. Outside the shadow of the detail to be observed let the screen brightness be H_2, in the middle of the shadow H_1 (see *fig. 3a*). At the edge of the shadow, due to the screen blurring, the brightness does not fall abruptly from the value H_2 to the value H_1, but there is a transition region of width O_s in which the brightness gradually varies from H_2 to H_1 (for the sake of simplicity we consider this variation to be linear). For the perceptibility of the detail drawn in fig. 3*a*, where the diameter d is larger then the screen blurring O_s, this gradual transition of brightness has no unfavourable result. If the contrast between H_1 and H_2 is equal to or larger than the minimum contrast which the observer can still just distinguish at the angle of vision corresponding to d, the detail will be observed in spite of the blurring [1]).

The situation becomes different when the detail is smaller than the screen blurring. In fig. 3*b* it may be seen that the regions of the transition in brightness at opposite edges now partially overlap. The brightness behind the detail nowhere falls to the low value H, *i.e.* the screen blurring results in a decrease in contrast. Thus if a detail of the size $d < O_s$ without the screen blurring has just sufficient contrast to be observed, then

[1]) With decreasing angle of vision the value of the minimum perceptible contrast increases. Due to the cutting down of the region of full contrast of a detail following from fig. 3*a* (difference in brightness H_2—H_1), therefore, perceptibility is somewhat unfavourably affected. The impression of contrast upon a gradual transition in brightness $H_2 \rightarrow H_1$ is also probably smaller than upon an abrupt transition.

owing to the decrease in contrast caused by the blurring it will remain below the required size. In this case we may call O_s the critical size of detail.

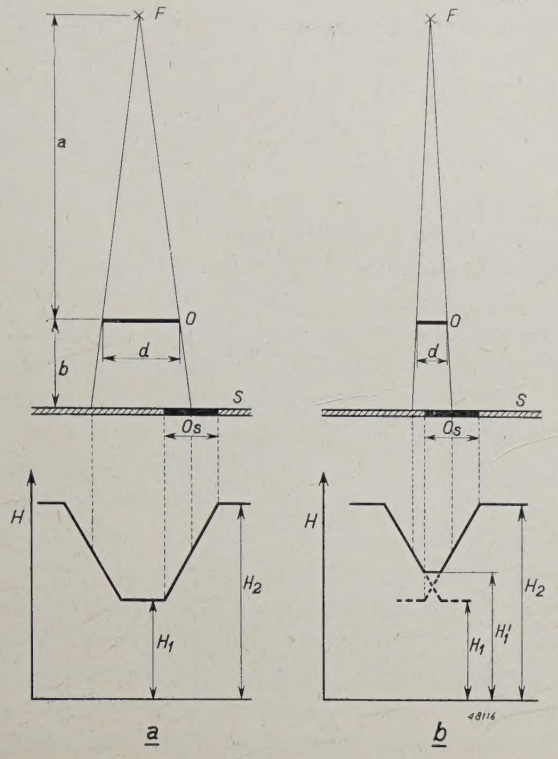


Fig. 3. Influence of the intrinsic screen blurring O_s on the projection of a detail of size d. The focus is assumed to be a point, so that the geometrical blurring $O_g = 0$. The distance b is assumed to be still so small that no appreciable enlargement occurs ($v \approx 1$). Below, the brightness H on the fluoroscope screen is plotted as a function of the position along a cross-section.
a) Large detail, $d > O_s$.
b) Small detail, $d < O_s$. The blurring here causes a decrease in contrast; $d = O_s$ in this case is the "critical detail size".

Let us now consider the same case when screening with magnification ($v > 1$). The shadow picture of the detail becomes v times as large, and the constant screen blurring O_s can now only cause a decrease in contrast when $v \cdot d < O_s$, *i.e.* in the case of details which are v times as small as before. Thus only v times smaller details are apt to escape observation. This is the effect with which we are concerned in the enlargement.

The effect is partially neutralized due to the fact that upon enlargement the total blurring is in any case increased; as already mentioned the geometrical blurring now comes to the fore. If we assume for the moment that only the geometrical blurring O_g is present ($O_s = 0$), then the situation for a large and a small detail is represented by *fig. 4a* and *b* respectively. O_g decreases the contrast for small details in exactly the same way as explained above for the screen blurring O_s. "Decrease of contrast" is now only another expression for the simple fact that the fluoroscope screen is no longer affected by the core shadow of the detail (see fig. 4*b*). The condition for decrease of the contrast is, analogous to the above, $v \cdot d < O_g$ and since

$$O_g = (v-1) \cdot f \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

(see equations (1) and (2)), the perceptibility is unfavourably affected for details with

$$d < f\,(1 - \frac{1}{v}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

If, therefore, the enlargement were very large ($v \gg 1$), so that the screen blurring becomes negligible compared with the geometrical blurring, then according to equation (4) the critical size of detail simply becomes equal to the width of focus. From this the importance of making the focus as small as possible for magnification technique becomes quite evident.

The simultaneous occurrence of geometrical and screen blurring results in a total blurring O_t, which can be approximately represented [2]) by

$$O_t = \sqrt{O_g^2 + O_s^2} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Decrease in contrast occurs at $v \cdot d < O_t$, and thus

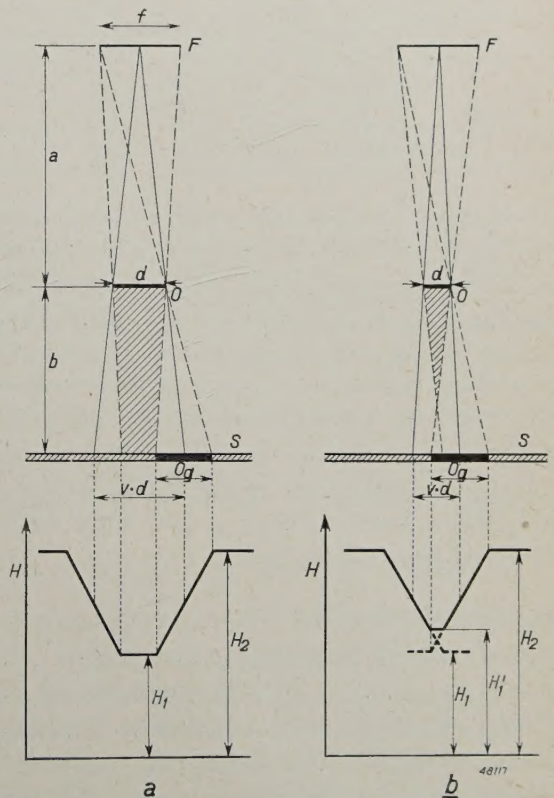


Fig. 4. Influence of the geometrical blurring O_g. It is assumed that $O_s = 0$. The figure is analogous to fig. 3.
a) Large detail, $d > O_g$.
b) Small detail, $d > O_g$. There is a decrease in contrast.

[2]) R. R. Newell, Brit. J. Radiology **30**, 493, 1938.

the critical size of detail in the general case is

$$d = \frac{1}{v}\sqrt{(v-1)^2 f^2 + O_s^2} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

or, if the formula is written in the non-dimensional form

$$\frac{d}{O_s} = \frac{1}{v}\sqrt{(v-1)^2\left(\frac{f}{O_s}\right)^2 + 1} \quad . \quad . \quad . \quad (7)$$

In *fig. 5* this relation is represented graphically. The ratio d/O_s is plotted as a function of the enlargement v for different values of the parameter

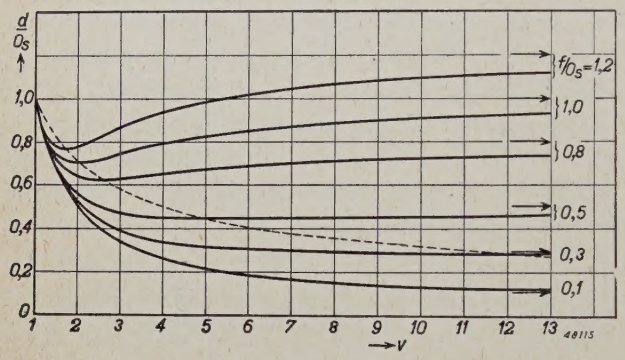


Fig. 5. In the presence of a screen blurring O_s and simultaneously a geometrical blurring (focus width f) the critical size of particle d, below which a decrease in contrast occurs due to the blurring, is given by equation (7). The ratio d/O_s is here plotted according to this equation as a function of the enlargement factor v, with the quotient f/O_s as parameter. Each curve passes through a minimum at an enlargement factor $v_m = 1 + (O_s/f)^2$; the level of these minima is given by $(d/O_s)_{min} = 1/2\sqrt{v_m}$ (dotted-line curve). With very small foci, for instance for $f/O_s = 0.3$, however, at an enlargement of $v \approx 4$ practically the whole theoretical gain is already obtained.

f/O_s. It may be seen that with a small enough focus the critical size of detail steadily decreases with increasing enlargement, but that at the same time this decrease is limited by the width of focus f. At the value $f = 0.3$ mm, which has been realized in the X-ray tubes to be described later, and with $O_s = 1.0$ mm, the practical enlargement will be not more than 4 times, since the further decrease of d is then no longer appreciable. The critical size of detail here is $d = 1/3$ mm, while without enlargement $d = O_s = 1$ mm!

The scattered radiation

In addition to diminishing the influence of the screen blurring enlargement also has another favourable effect which lies quite outside the considerations already discussed. The longer the distance (b) between object and fluoroscope screen, the smaller the percentage of secondary X-radiation [3]) scattered by the object which falls on the screen. This scattered radiation gives a uniform fog over the image which reduces all contrasts — an effect, therefore, which in the case of enlargement becomes less potent or even almost disappears. A gain in contrast is thus obtained over the whole image. The degree to which the contrasts may increase is illustrated by the fact for instance in lung fluoroscopy the scattered radiation may have the same intensity as the primary radiation and in screening of the abdomen its intensity may even be three or four times as great [3]).

The effect of the enhanced contrasts is further increased by the fact that upon enlargement the visual angles within which all the details of the object are seen increase proportionally. For details with a larger visual angle an observer has a higher sensitivity to contrast (*cf.* footnote [1])).

In *fig. 6* two photographs (*b* and *c*) are given of an object (*a*) on the fluoroscope screen with and without enlargement. It may be seen that small details originally invisible or almost so are rendered visible and the contrasts in the enlarged image become larger. For the sake of comparison a reproduction (*d*) is also given of the normal fluoroscopic image subsequently enlarged optically. Since the screen blurring is here also enlarged and the fog due to the scattered radiation remains, no improvement is obtained.

Ease of observation

A comparison of figures 6*b* and 6*c* also gives an idea of the more convenient and easier observation of the enlarged image to which reference has already been made. The concrete advantage of this in practice is that perception of details is much quicker and less fatiguing.

The screen brightness

Influence on the improvement obtained

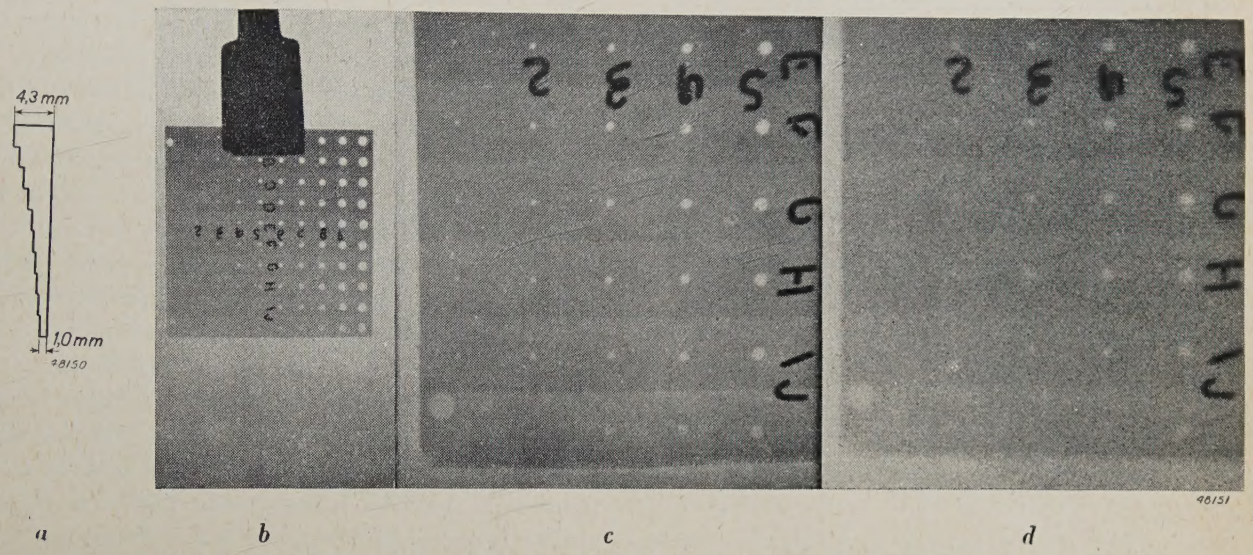
It would be premature to conclude from the above that in every case (with sufficiently small focus) enlargement would be an advantage. We have untill now disregarded the fact that upon enlargement the brightness of the image on the fluoroscope screen may decrease.

We must here make a distinction between the contribution to the brightness of the primary (the image-forming) and that of the secondary (scattered) X-radiation. The contribution to the brightness of the secondary radiation decreases in any case upon enlargement. This means that all kinds of physiological factors which remain to be discussed are affected unfavourably to some extent, but on

[3]) See for example W. J. Oosterkamp, Combatting the scattered radiation in the medical X-ray image, Philips Techn. Rev. 8, 183, 1946.

the whole the disappearance of the scattered radiation can only be favourable. However, the contribution to the brightness of the primary radiation may also decrease appreciably upon enlargement. We shall see presently that in the testing of materials this is very often the case, while in medical diagnostics it need not occur invariably.

With decreasing brightness of screen the visual acuity and sensitivity to contrast of the observer diminish. The advantage of the increase in all contrasts by enlargement (including the already mentioned higher sensitivity due to the larger angles of vision) is partly neutralized by a diminution of the sensitivity to contrast. As to the blurring the critical size of detail caused thereby does not depend upon the brightness, but the decrease in contrast in the case of details with dimensions smaller than the critical, when the observer's sensitivity to contrast decreases, becomes fatal already at greater contrasts, so that a larger group of details falls into the category of those no longer perceptible. Part of the gain which enlargement brings especially for the observation of small details poor in contrast is thus lost again.



a *b* *c* *d*

Fig. 6. Normal (*b*) and four-fold enlarged (*c*) fluoroscopic picture of a test object (*a*) consisting of an aluminium plate with stepwise varying thickness and in each step a series of holes of different diameter. *d*) The same picture as in (*b*), *i.e.* obtained with normal fluoroscopy, but reproduced on a scale four times as large as (*b*); no appreciable improvement compared with (*b*). The fact that in the reproductions of this review more details can be discerned in *d* then in *b* is due to the new „blurring" which the lattice of the autotype introduces; in this lattice small details disappear rather than large ones.

In addition to the (small and large) details with poor contrast to which all the above discussion refers, we must also pay attention to small details with high contrast. In addition to the advantage of the sharper reproduction on the whole, enlargement gives for these details no improvement in perceptibility. A loss in brightness which accompanies the enlargement may even worsen the perceptibility of such details. This may be explained as follows:

The perceptibility of the details rich in contrast is limited by visual acuity, *i.e.* by the minimum angle of vision α_{min}, in which the observer can still just distinguish a detail. When the brightness decreases due to enlargement, α_{min} increases. At the same time, however, the angle of vision α of the detail increases proportionally with the enlargement factor v. The perceptibility is unfavourably affected when α_{min} increases more rapidly than v. The relation between α_{min} and the brightness is known from a number of investigations. The relation between the brightness and v may, as already indicated be very different. If we assume, for example, a variation of the screen brightness proportional to $1/v^2$, it is then found that as long as the brightness remains above a certain limit the increase of α remains ahead, so that upon enlargement the "effective" visual acuity improves. At what enlargement this limit of brightness is met depends, of course, on the initial brightness (for $v = 1$) and this in turn is quite different according to the nature and thickness of the object. Enlargement above a certain factor v, which also depends upon the object, will thus have an unfavourable effect on the observation of the very small details rich in contrast, although the visibility of the details poor in contrast is improved. This fact may be of importance in the testing of materials.

Screen brightness and method of enlargement

Beginning with normal fluoroscopy where object and screen are set up close to each other, it is possible to enlarge in two ways:

1) the object and the X-ray tube can be left in the same positions and the screen placed farther away (a constant, b and thus also $a + b$ larger);
2) the screen and the focus can be left in position and the object placed closer to the focus (a smaller, b larger, $a + b$ constant). The screen

brightness is inversely proportional to the square of the distance $(a + b)$ between focus and screen; in the first method, therefore, upon enlargement a considerable decrease in brightness is obtained, while in the second method the (primary) brightness remains constant.

In order to avoid the above-mentioned disadvantages of a decrease in brightness, the second method seems to be the better one. For the testing of materials, however, it cannot always be applied; in particular it cannot be applied in those cases where with normal fluoroscopy, in order to obtain the greatest possible screen brightness, the object and the screen are placed as close to the X-ray tube as the dimensions of the object allow. It is therefore only possible to enlarge in such a case by placing the screen farther away, and the decrease in brightness proportional to $(a + b)^2$ or, according to equation (2), proportional to v^2, must be accepted into the bargain.

In medical diagnostics, on the other hand, with normal fluoroscopy the patient is always placed at a fairly great distance from the focus in order to obtain an overall picture of the part of the body concerned [4]). Increasing the screen brightness by placing patient and screen closer to the focus would be of no use here since a greater brightness could also be realized by increasing the current through the X-ray tube. The fact that this is not done and thus that the focus is not loaded to anywhere near the permissible limit is due to the limitation of the X-ray dose which may be administered to a patient in fluoroscopy. The intensity of radiation on the patient, by which the screen brightness is also approximately determined, is chosen so low that the doctor may devote several minutes to the perception of the fluoroscope image and still remain far below the permissible dose. If it is now desired to pass over to enlarged fluoroscopy the second method can in this case be applied, *i.e.* the patient can be placed closer to the focus provided a correspondingly shorter time of observation is allowed in order to keep the dose on the patient within safe limits in spite of the stronger intensity of radiation. Since the distance between focus and screen here remains unchanged the (primary) screen brightness with enlarged fluoroscopy is no less than with normal screening.

Screen brightness and focus loading

We have explained that for enlarged screening a much smaller focus is necessary than is normally used. If the focus is normally loaded to the limit of the power it can receive continuously per unit of surface without becoming too hot, decrease in size of the focus necessarily means a reduction in the current in the X-ray tube and thus in turn a decrease in the screen brightness. In testing materials where powers of 2 kW, for instance 150 kV_{max}, 20 mA and even more are used continuously, with the normal large foci the permissible focus temperature is indeed approached. Nevertheless, the decrease in size of the focus need not have such an unfavourable effect on the screen brightness as would follow from the ratios of the surfaces of the foci. A smaller focus has a higher specific loading capacity (at the same focus temperature). This is due to the fact that the heat developed on the small focus is dissipated not only towards the interior of the anode but to an appreciable extent also laterally (edge effect). Moreover, the specific focus loading capacity can be further increased by the employment of a rotating anode [5]).

In medical diagnostics, due to the above-mentioned limitation of the dose on the patient, with normal fluoroscopy the permissible specific focus loading is far from being attained. A tube voltage of for instance 75 kV_{max} with a current of 2 mA is used, *i.e.* a continuous loading of about 150 watts. It is here unnecessary to have the anode rotate. If the focus is now decreased in size, the focus temperature increases and at the smallest foci realized, of 0.3 mm, it is found necessary to let the anode rotate in order to be able to employ the above-mentioned 150 watts [6]).

Constructions of X-ray tubes with very small focus

The requirement for obtaining a small focus is a strong concentration of the electron beam emitted by the filament of the X-ray tube. In tubes for testing materials, where the focus must sometimes

[4]) If the patient is close to the focus the relative differences in distance for various details in the body are so large that some are projected considerably larger than others.

[5]) See for example J. H. van der Tuuk, X-ray Tubes with Rotating Anode ("Rotalix" tubes), Philips Techn. Rev. **8**, 33, 1946; also Philips Techn. Rev. **3**, 296, 1938.

[6]) It should be pointed out that the enlarging technique, at least as far as the improvement in definition is concerned, does not as yet offer advantages in making X-ray photographs. It is true that here too there is a kind of "screen blurring", namely that caused by the grains of the usual reinforcing screens. But in enlarged photography in the case of moving objects the blurring due to motion would increase very much, since the small focus for instantaneous photography in any case can be less heavily loaded than a large focus and for that reason alone the exposure would have to be longer. For stationary objects there is no objection to longer exposure, but just for that reason it is then possible, if greater definition is desired, to omit the reinforcing screens, so that only the very slight intrinsic lack of definition of the X-ray film itself remains.

be introduced into all kinds of cavities, the anode with the focus is often placed in a projection of the tube at a distance of 20 to 30 cm from the filament. For the small as well as for normal foci the required concentration can then only be attained by focusing with the help of one or more magnetic coils. This method gives good results but is rather elaborate. With a given magnetic field the focusing is only good for a given velocity of electrons. If it is desired to vary the voltage on the X-ray tube according to the object to be examined, the exciting current of the magnetic coils must also be accurately adjusted anew each time. If an extremely small focus is desired, this adjustment has not only to be very precise but, moreover, for the proper functioning of the tube a carefully smoothed and constant D.C. voltage is necessary, since otherwise the size of focus alternates appreciably with the ripple of the tube voltage.

In X-ray tubes where cathode and anode are placed only about 1 cm apart — in medical diagnostics such tubes are used exclusively — the problem can be solved much more satisfactorily by employing electrostatic focusing. This means that by giving a suitable shape to the cathode the potential between cathode and anode is made to vary in such a way that a strong electrostatic "lens" is formed. This focusing is independant of the tube voltage. The potential of every point in the deflecting field changes proportionally with the tube voltage; since the deflecting action on an electron is proportional to this potential, and on the other hand inversely proportional to the tube voltage which accelerates the electron, the two effects exactly cancel out each other.

The tubes with very small focus developed in the Philips laboratories (small focus tubes") employ electrostatic focusing. In fig. 4 the cathode of such a tube is shown. It may be seen that two spiral filaments are assembled in grooves of the cathode, one large and one very small. These can be switched on according to choice. The thin spiral furnishes the small focus of 0.3 mm width, the thick one gives a focus of 2 mm or, in other tubes, of 1.2 mm, which is intended for normal fluoroscopy and for making photographs.

The size of the focus obtained is checked in the familiar way by making a photograph of it with a pinhole camera. The aperture of this camera, which is normally made 0.2 mm, must in this case be considerably smaller (about 0.03 mm), in order to obtain a sufficiently sharp projection of the small focus.

Since the tube also serves for taking X-ray photographs when the larger focus is used, the anode is of a rotating type which in any case is desirable in order to be able to give the small focus a high load. In the usual constructions of rotating anodes

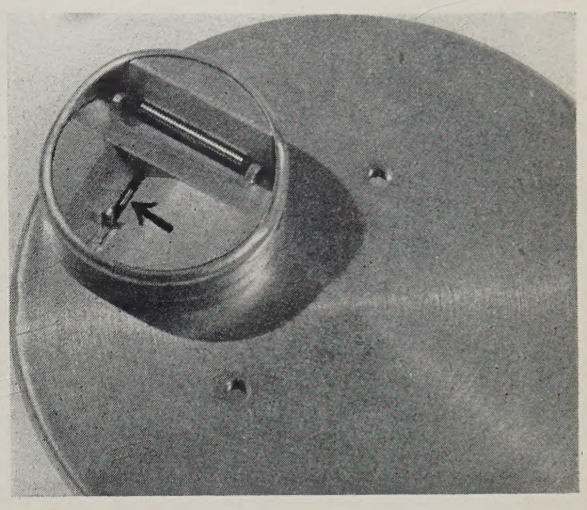


Fig. 7. Cathode of a double-focus X-ray tube, with rotating anode, specially developed by Philips in connection with the enlargement technique. The thick filament gives a focus of 2.0 mm width, with which photographs are made, the thin filament gives a focus of only 0.3 mm width, intended for normal and enlarged fluoroscopy. Thanks to the special construction of the cathode cap the electrons emitted are focused electrostatically on the anode.

the rotor runs in two ball-bearings. The guiding rings of the latter in the rotor as well as on the anode carrier have several tenths of a millimeter axial play in order to take up the difference in thermal expansion of rotor and anode carrier. For working with the small focus, however, the play referred to is undesirable. Axial vibrations of the anode might be caused; since the focus is thereby displaced perpendicular to the direction of the effective beam of X-ray, a relatively considerable increase in the apparent width of focus occurs. In order to avoid this effect a special bearing of the rotor has been made for which the axial tolerance is limited to several hundredths of a millimeter.

Practical application of enlarged fluoroscopy

In the testing of materials

The above statements about the limitation of the dosage in medical diagnostics might give the impression that the general situation for normal fluoroscopy is much more favourable for testing materials. It is possible to work with a very small distance between focus and object, with the maximum focus loading, with arbitrarily long times of observation (apart from economical considerations). These advantages, however, are practically neutralized by the disadvantage that in testing mater-

ials the objects dealt with have a much greater absorption. An iron plate 0.02 mm thick or an aluminium plate of 4.5 mm absorbs (with the same tube voltage of 75 kV) just as much as the chest of a patient. Thus in material testing the screen brightness is by no means so much more than sufficient as might be supposed. Moreover, in a workshop it is impossible to use such low screen brightness as can be used in the X-ray room of a medical establishment. The result is that the employment of fluoroscopy in material testing has always been limited to the light metals with relatively little power of absorption, such as aluminium and magnesium alloys.

Enlarged screening in the testing of materials, as was explained above, involves opposing effect in definition and brightness. Such is also the case, however, with normal screening technique. The screen blurring of 1.0 mm already mentioned is only valid for very sensitive screens having a high brightness. If a less sensitive screen were used, which gave for instance only half the brightness, the blurring would only be half as great. The compromise, which in normal fluoroscopy leads to the choice of a given screen, must be considered anew in enlarged fluoroscopy, and it is thus *a priori* still a question whether enlargement is really an advantage in material testing. Experience has answered this question in the affirmative and in the screening of light metals magnification technique has been employed for a number of years. We shall not go more deeply into the results obtained but refer only to the literature [7]).

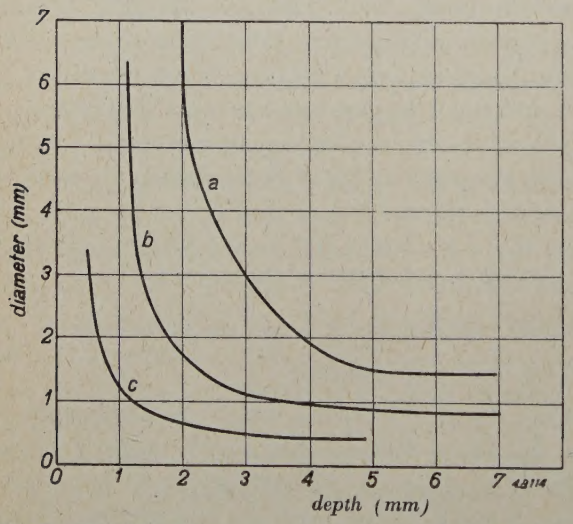


Fig. 8. Visibility tests on a phantom of "Philite" which is about equivalent to the human lung for X-ray fluoroscopy. Holes of different depth and diameter were bored in the phantom. For every depth (in mm) the diameter of the smallest perceptible hole (also in mm) is plotted.
a) For fluoroscopy without enlargement, focus width 0.4 mm, tube voltage 70 kV_{min}, tube current 1.4 mA.
b) For fluoroscopy with four-fold enlargement; data as in (*a*)
c) On an X-ray photograph under optimum conditions.

In medical diagnostics

We have seen that in medical diagnostics enlarged

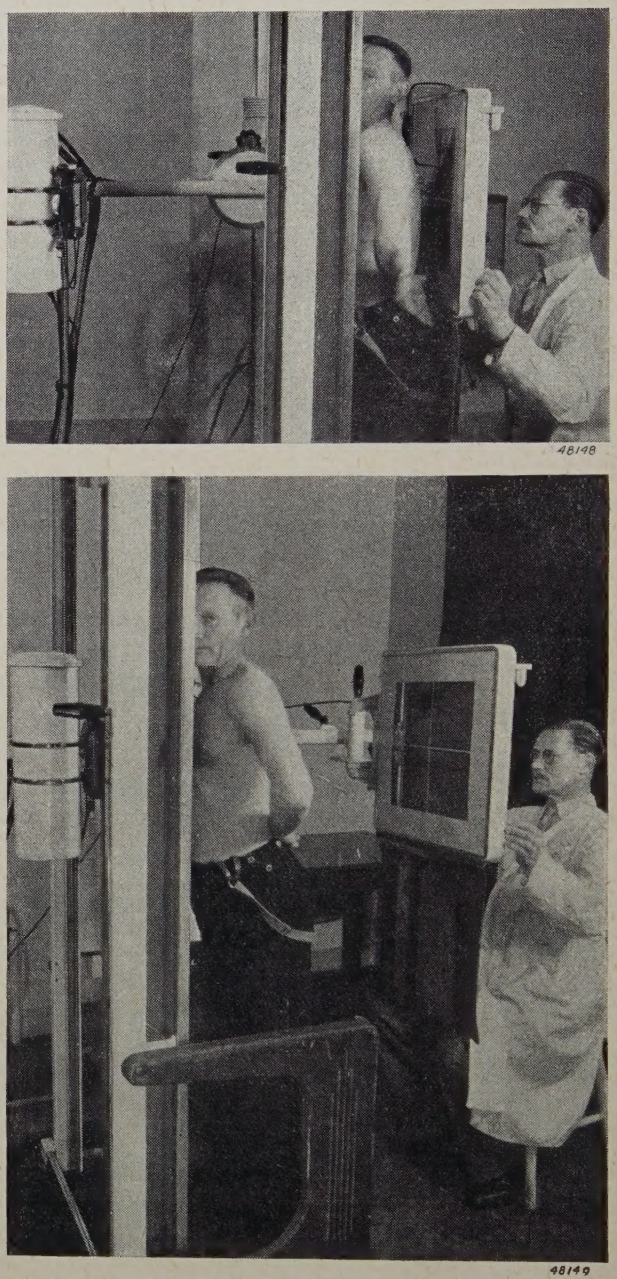


Fig. 9. Improvised arrangement for the enlarged screening of lung patients at the consultation bureau of the Medical Department of the Philips Factories. The patient's position is fixed, the fluoroscope screen and the X-ray tube are rigidly connected with each other and can be moved by the practitioner. Since in the periodically repeated, normal fluoroscopy the patient always stands with his back to the tube and thus the surface of the back always receives the highest radiation intensity (four times as much as a point in the middle of the lungs), for greater safety the patient is here placed with his chest facing the tube.
a) Situation for normal fluoroscopy.
b) Situation for enlarged fluoroscopy.

[7]) See for example R. Berthold, Atlas der zerstörungsfreien Prüfverfahren T/Rö 10/7-10, J. A. Barth, Leipzig 1938.

fluoroscopy need not lead to a loss in (primary) screen brightness, but only to a limitation of the time of observation. The favourable results which could be expected from this led to a systematic investigation of the possibility of using enlarged fluoroscopy in the Medical Department of the Philips factories.

This investigation, of which several preliminary results have been published elsewhere [8]), was mainly in two directions. In the first place observation tests were performed with a phantom of "Philite" which, as far as the absorption and scattering of X-ray is concerned, corresponds approximately to the chest of a patient. A number of holes of different diameter and depth were bored in the phantom. The graphic representation in *fig. 8* shows the visibility of these holes with normal fluoroscopy (*a*), when screening with four-fold enlargement (*b*) and on a good X-ray photograph (*c*). It may be seen that there is a significant improvement between normal and enlarged fluoroscopy. Moreover, these experiments already showed how important it was that the perception could take place more easily and more quickly with the large image than with the small one. Because of this the necessary shortening of the time of observation of a patient (to for instance 10 to 20 sec.) already mentioned is less of a disadvantage.

In the second place a group of lung patients in the routine examinations of the consultation bureau were examined not only with ordinary fluoroscopy but also with enlarged image. The improvised arrangement with which these experiments were carried out is shown in *fig. 9*. A standard on rollers supports the X-ray tube and the fluoroscopy screen. The screen first lies close against the patient so that a normal orientation image is obtained. After having examined this briefly and determined what part of the image requires special attention, the observer takes one or two steps to the rear drawing with him the standard with screen and tube, while the patient remains in his place. The observer can now see the part in question more clearly on the enlarged image.

One of the disadvantages for the doctor is that upon enlargement he literally loses contact with the patient. In normal fluoroscopy it is very useful to be able to turn the patient slightly and to move him about so as to make use also of the sense of touch.

When using enlargement this is not possible, because the time is too short and, moreover, the doctors' arms are not long enough to reach the patient from his position behind the more distant screen.

In spite of this the investigation gave encouraging results. From statistics of the small number of 51 cases examined [9]) it was found that by enlarged fluoroscopy in the case of 10 patients better insight was obtained into the structure of a lung process, 8 times the presence or absence of a cavity was ascertained with greater security (the diagnosis was later checked with the still better photographic image) and in general in 35 cases an improvement was noted, compared with 16 cases where enlarged fluoroscopy gave no concrete advantage.

One of the practical problems still to be solved before enlarged fluoroscopy can become part of the daily practice of the röntgenologist is the adequate protection of the patient against too high doses of X-radiation. This will be more necessary, because just the difficult or interesting cases which form the most fruitful objects for enlargement technique are usually observed for a longer time even in normal fluoroscopy. And this holds even more for the screening of the abdominal organs than for that of the lungs, since in the former case, due to the greater absorption, higher radiation intensities and in some cases (stomach examinatons) longer times of observation must be employed.

[8]) G. C. E. Burger, Hand. 29 Ned. Nat. Gen. Congres, Amsterdam 1943, p. 47. Fig. 8 is also borrowed from this.

[9]) G. C. E. Burger, not yet published.

THE TECHNIQUE OF INVESTIGATIONS WITH RADIOACTIVE AND STABLE ISOTOPES

by A. H. W. ATEN Jr. and F. A. HEYN. 539.155.2:539.167.3:539.16.08

When tracers (tagged atoms) are used for scientific or technical investigation, one is concerned with radioactive and, in some cases, with stable isotopes of certain elements. In as far as the choice is not prescribed by the problem itself, it is often the half-value time of the radioactive isotopes which determines it; the half-value times must be neither too long nor too short. Only in the cases of oxygen and nitrogen is one now limited to the use of stable isotopes, while heavy hydrogen is also still much used. The preparation of the radioactive isotopes is by means of nuclear reaction brought about, for example, with the help of a cyclotron, a high-voltage apparatus or in an uranium pile. Stable isotopes are obtained by the separation of the mixtures of isotopes occurring in nature, for instance by fractional electrolysis, distilation or diffusion. The synthesis of possibly desired compounds of tagged atoms is in practice still one of the most difficult problems that can occur in the application of the method. A neat solution is sometimes found by a biological method of synthesis, *i.e.* in micro-organisms or in the animal body. The analysis of the mixtures of isotopes obtained when applying this method is carried out, in the case of the radioactive isotopes, by measurement of the radioactivity, for instance with an electron counter; many details of this are discussed in this article. For accurate localization of the radioactive isotopes, auto-radiograms are very well suited. Mixtures of stable isotopes are analysed with the mass spectrograph, or, as in the case of heavy hydrogen for instance, by measurement of density.

The concise survey of the tracer or "tagged" atom method and its possibilities of application [1]) which we gave in the previous number of this periodical will be supplemented here by a discussion of the manner in which the method is applied.

The following is a resumé of the most important points in the article mentioned.

Of every chemical element occupying a certain position in the periodic system there are different kinds of atoms, differing in atomic weight. These mutually "isotopic" chemically indistinguishable kinds of atoms may be either stable or radioactive. When two isotopes of an element take part in a physical or chemical process they behave in practically the same way. Nevertheless one isotope, either because it is radioactive or because of its different atomic weight, can in principle be recognized, localized and quantitatively determined by the investigator at every stage of the process. These atoms are, as it were, tagged and may be used as "tracers".

Use is made of this fact for the investigation of very diverse processes, firstly because of the great eare with which the place, quantity and distribution of a radioactive substance is determined, and because of the extremely great sensitivity which can easily be attained; and secondly because certain types of processes have only been made accessible for investigation by the tracer method, namely those processes where there is an exchange of chemically identical particles. Technical examples are the diffusion of a substance in an identical environment and the transfer of material upon friction between two identical surfaces. Especially in the field of chemistry and in the physiology of plant and animal life such processes play an important part.

If it is desired to carry out an investigation with the help of tagged atoms, the following are the main practical problems encountered.

1) What kind of atom should be chosen?
2) How is it obtained.
3) In what form (chemical compound) should the atoms be used and how can they be brought into that form?
4) How are the mixtures of isotopes occurring in the experiments analyzed?

The answers to questions 2, 3 and 4 will often be simple: the tagged atoms, or in some cases the desired compounds of them, are ordered from the nearest laboratory equipped for making them, and the mixtures to be analyzed are sent to the same address for analysis. The fact that various laboratories possessing the necessary equipment — in Europe, among others, the Philips Laboratory in Eindhoven — have opened the opportunity for the application of this method, has very greatly promoted the rapid spread of its use. However, even for those who are thus relieved of the difficulties involved in questions 2, 3 and 4, it will also be of importance to be informed abut the technique followed in solving these questions.

Choice of the kind of atom

In the first experiments with tracers carried out by Hevesy and others a radioactive isotope of lead was usually employed, and sometimes also thallium or bismuth. For example, the assimilation of lead ions by plants was investigated, or the distribution of lead administered between a tumor and the normal tissue in the animal body. Since the plant and the animal body normally

[1]) Philips Techn. Rev. 8, 296, 1946.

contain no lead, it would in principle also have been possible to carry out these experiments by means of a chemical analysis, but then there would have been the practical difficulty of determining the extremely small quantities of lead involved, for which the measurement of radioactivity is particularly suited.

It was obvious that such experiments, performed with substances which do occur in the plant or animal and which are involved in their natural functions, could furnish much more important results. Moreover, from the point of view of methodics, such experiments were more interesting, since chemical analysis must in principle fail here, and only the use of tagged atoms which are also recognisable in a chemically identical environment was able to make these experiments possible.

The reason why the pioneers of the method chose lead, thallium or bismuth was simply because at that time no other tagged atoms were available. The three mentioned are the only non-radioactive elements of which radioactive isotopes with sufficiently strong radio-activity occur in nature.

The situation is now quite different. Since the discovery of artificial radioactivity in 1934 we have learned how to make one or even several radioactive isotopes of every element. The development of methods for the separation of isotopes has also made it possible to prepare stable isotopes of a series of elements in sufficient quantities for practical use. There is such a wide choice of tagged atoms that one might almost speak of an "embarras de choix", were it not that there are again a number of factors limiting the choice.

To begin with, the choice is often prescribed by the problem itself. If, for example, it is a question of the diffusion of certain alloy components in steel and their influence on the properties of the alloy, because of the usually very specific influence of an alloying element, tagged atoms of that same element will have to be chosen. The same is the case with specific biological effects of certain elements, such as the assimilation of iron by anaemics.

An example of a problem permitting a certain freedom of choice is the previously described investigation of the friction between metal surfaces, at least in so far as it is only a question of the mechanism of the friction and not of possible specific effects with certain metals.

In many cases the choice is limited by the half-value time of the radioactive isotopes which can be used (*i.e.* the time in which the radioactivity decreases to one half of its initial value). A list of the half-value times of a series of often used isotopes is given in the table. The half-value time must be neither too short nor too long. If it is very short it means that the radioactivity of a sample decreases rapidly and, in spite of the great sensitivity of the measuring instruments, falls below the limit of detection so quickly that there is no time to perform the experiment in a proper way. This is especially true for biological problems, where, in contrast to chemical problems, the processes to be investigated cannot be accelerated by raising the temperature or otherwise. If the half-value time is very long it means that only a few atoms disintegrate per unit of time, the radioactivity then being so slight already from the beginning that it may not be possible to measure it.

The curves in *figs. 1a* and *b* illustrate these facts. They also show that both with too short and with too long half-value times conditions for the measurements are improved when the concentration of the radioactive isotope in the substance is sufficiently increased. For instance, it was originally very difficult to carry out experiments with radioactive carbon, while because of the dominating position of carbon in all organic chemical compounds

Table I. Half-value times of several radioactive isotopes *)

The symbol of each chemical element is accompanied by an index (upper right) indicating its so-called mass number (atomic weight in whole numbers) and another (lower left) giving its number in the periodic system (nuclear charge).

Element	Atomic weight of the isotope mixture occurring in nature	Symbol of the radioactive isotope	Half-value time
Hydrogen	1.0081	${}_1H^3$	31 years
Carbon	12.010	${}_6C^{11}$	20.5 minutes
		${}_6C^{14}$	>1000 years
Nitrogen	14.008	${}_7N^{13}$	9.93 minutes
Oxygen	16.000	${}_8O^{15}$	2.1 minutes
Sodium	22.997	${}_{11}Na^{22}$	3.0 years
		${}_{11}Na^{24}$	14.8 hours
Magnesium	24.32	${}_{12}Mg^{27}$	10.2 minutes
Phosphorus	30.98	${}_{15}P^{32}$	14.3 days
Sulphur	32.06	${}_{16}S^{35}$	87.1 days
Chlorine	35.457	${}_{17}C^{38}$	37 minutes
Potassium	39.096	${}_{19}K^{42}$	12.4 hours
Calcium	40.08	${}_{20}C^{45}$	180 days
Manganese	54.93	${}_{25}M^{52}$	6.5 days
Iron	55.84	${}_{26}Fe^{59}$	47 days
Copper	63.57	${}_{29}Cu^{64}$	12.8 hours
Bromine	79.916	${}_{35}Br^{82}$	34 hours
Strontium	87.63	${}_{38}Sr^{89}$	55 days
Iodine	126.92	${}_{53}I^{126}$	13 days
		${}_{53}I^{128}$	25 minutes

*) Of most of the elements listed here several other isotopes exist which are not listed. The data are taken from G. Seaborg, Rev. mod. Phys. **16**, 1, 1944.

it was desirable to be able to carry out just such experiments. The isotope C^{14} with a half-value time of about 10 000 years was not active enough, and in the case of C^{11} with a half-value time of only 21 minutes the experiments often had to be given up

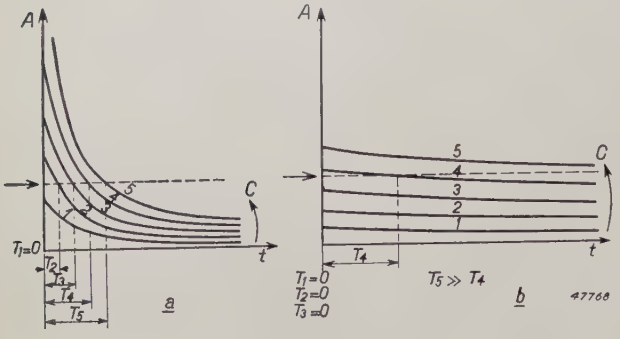


Fig. 1. Disintegration curves of a radioactive mixture of isotopes *a*) with a short half-value time, *b*) with a long half-value time. Each curve holds for a definite initial concentration of the radio-active isotope in the mixture, the concentration increasing in the direction of the arrow. The dotted line indicates the minimum activity that can be detected with the measuring instruments. T_1-T_5 are the respective times during which a measurement of the activity is still possible with the five preparations in question. The limit of the measurement indicated by the dotted line will in general be lower for substances with long half-value times than for those with short half-value times. In the first case the measurement of the activity can be accomplished by counting current impulses for a longer time, so that with fewer impulses per unit of time the statistical fluctuations of activity and zero effect are sufficiently exceeded (see under fig. 2 and at the end of the article).

prematurely. Later on, however, C^{11} was successfully prepared in such large concentrations that experiments concerned with the photosynthesis of the carbohydrates in plants could be continued with it for more than 5 hours. This corresponds to a decrease in the activity to about 1/30 000 of the initial value; actually the difference in activity between initial and final substance may be very much greater still, since the nature of the experiment may determine the fact that the final substance contains only a fraction of the original amount of carbon.

If it is not possible in practice to realize the required initial concentration of an isotope with an unfavourable half-value time, there is still the possibility of using a stable isotope of the element in question, even though it is usually much more difficult to work with than the radioactive isotopes. It is unfortunate that the available radioactive isotopes of just the most important elements in organic chemistry — carbon, hydrogen, oxygen and nitrogen — have decidedly unfavourable half-value times. In the case of nitrogen and oxygen, for most experiments one must still use the stable isotopes N^{15} and O^{18}. For experiments with tagged hydrogen, the heavy hydrogen isotope with atomic weight 2, deuterium (H^2 or D), was formerly used on a large scale and is still being used.

In the case of a high initial activity, which according to fig. 1*a* is needed in the case of isotopes with a short half-value time, there may be a complication. When metabolic experiments are being performed on animals it must be taken into account that the radioactive radiation not only acts as a "spy" but may also have a destructive effect on tissues of the test animal. There is fortunately a tolerance dose which must be exceeded before any harmful effect of the radiation is manifested. If the initial activity is so great that this dose is exceeded, care must be taken that the progress of the experiment is not altered as a result of the irradiation.

In this connection it should be pointed out that it is one of the great advantages of the tracer method that, in principle, the metabolism can be studied without disturbing the processes thereof in normal living animals.

Incidentally we may mention here the remarkable fact that not only radioactive but also stable isotopes in too high concentrations can damage the life functions of a test animal. This was ascertained in experiments on the assimilation and excretion of water in test animals. The animals were allowed to drink water of which the molecules to a known percentage were built up of heavy instead of ordinary hydrogen, and the density of the water was then studied in the blood, lymph, urine, etc. In the human body, which consists for 80% of water, it was found that half of all the water in the body was renewed in 13 days. In goldfishes, which were simply allowed to swim in "heavy water", it was found that the water molecules in a fish are practically entirely renewed within a few hours. When, however, the concentration of D_2O in the water was very high, so that many H_2O molecules in the fish were replaced by D_2O, this caused the fish to die.

The preparation of tagged atoms

The radioactive isotopes not occurring in nature can be formed by certain transmutations of atomic nuclei (nuclear reactions). The following are examples of such transmutations:

$$_{16}S^{32} + {}_0n^1 \rightarrow {}_{15}P^{32} + {}_1H^1,$$
$$_{15}P^{31} + {}_1H^2 \rightarrow {}_{15}P^{32} + {}_1H^1.$$

In the first case the unstable phosphorus isotope P^{32} is formed from a sulphur atom and a neutron, in the second case from ordinary phosphorus and deuterium. These nuclear reactions do not take place, like chemical reactions, as soon as the electron clouds of the atoms approach each other, for which only little energy is necessary, but they require the direct influence of certain particles (for instance protons or neutrons) on the nuclei, which influence has to be brought about by causing the two to collide with each other with very great force. In our first example the necessary neutrons are obtained by shooting deuterium ions at a preparation containing lithium, the neutrons thus liberated being allowed to enter a vessel containing carbon disul-

phide in which, according to the equation, sulphur nuclei are transmuted into $_{15}P^{32}$. After a certain amount of ordinary phosphorus has been added in order to obtain quantities which can be dealt with practically, this phosphorus, together with the ordinary phosphorus, can be isolated by chemical means.

If the final product of a nuclear reaction is an isotope of the original atom, as in our second example, a chemical separation would seem impossible at first glance. If, however, one begins with a compound of the element to be transmuted, each atom transmuted will be thrown out of the molecule by the shock, so that the reaction product is in the form of free atoms or ions, which can then be isolated from the compound.

The strong forces required for nuclear reactions can be given to the atoms in a cyclotron, in which the (ionized) atoms are accelerated along a spiral trajectory to enormous velocities corresponding to voltages of several million or even tens of millions of volts; or in a high-tension apparatus working with a voltage of several hundred thousand or millions of volts [2]). The efficiency of the nuclear reactions, however, is generally very low, so that very intensive beams of bombarding particles are necessary to obtain the desired concentrations of tagged atoms. The apparatus for the preparation of the tagged atoms therefore becomes so elaborate and expensive that it can only be installed in a few laboratories. As already mentioned, however, various institutes are already making available to every investigator the preparations of tagged atoms which they make. It should be pointed out that a new possibility for obtaining radioactive elements has been opened by the construction of the uranium piles which played such an important part in the development of the atom bomb; numerous radioactive substances can be prepared with their help.

The separation of the radioactive substance formed in the nuclear reaction is not necessary in every case. The experiments already mentioned on the friction between metal surfaces, for example, were carried out with metal plates which had been irradiated in a cyclotron, whereby a small part of the atoms at the surface were transmuted *in situ*. The plates could be used without further treatment.

A similar case is found in a certain method of chemical analysis. If in a mixture of different elements one element occurs in which radioactivity can be excited by irradiation, without a radioactivity of comparable half-value time being excited in any of the other elements, it is possible to conclude from the intensity of the radiation whether, and in what concentration, that element is present. In this way it was possible to demonstrate in a simple way the presence in paper of traces of sulphur which could not possibly be detected by the methods of analytical chemistry, by exposing the paper to the intense neutron radiation of a cyclotron and afterwards demonstrating the presence of a radioactive transmutation product of sulphur. The tracer method here appears in a rudimentary form: with the preparation and detection of the tagged atoms their role is finished, and they need no longer take part in any process.

We can be brief about the preparation of stable isotopes as tracers. Of the elements occurring in nature many consist of different stable isotopes which are mixed in fixed proportions. Nitrogen, for example, consists for 99.6% of N^{14} and for 0.4% of N^{15}. These components can be separated by making use of the difference in velocity of diffusion or reaction, which differences are based upon the difference in atomic weight. Although when the mixture is subjected for instance to a diffusion process the ratio between the isotopes at the end of the process will differ by only a few percent from that at the beginning [3]), if the mixture thus obtained is again subjected to the same process and this is repeated many times, very large deviations from the original ratio of isotopes can finally result. This will occur more quickly the greater the relative difference in mass, thus the lighter the element. Heavy water, which can be concentrated by the electrolysis of water, since the normal hydrogen disappears from the water at the cathode more quickly than the heavy hydrogen (originally present in a ratio of 1 : 5 000 to the normal hydrogen), is obtained in a purity of 96% or higher after only a few steps.

The stable isotopes $_8O^{18}$, $_7N^{15}$, and heavy water are at present obtainable commercially. Heavy water was previously very expensive: a preparation with 96% D_2O cost several times as much as its own weight in gold. The situation has improved very much because of the fact that now, for instance

[2]) See for example F. A. Heyn and A. Bouwers, An Apparatus for the Transmutation of Atomic Nuclei, Philips Techn. Rev. **6**, 46, 1941. On the cyclotron see for instance W. B. Mann, The Cyclotron, Methuen's Monographs, London 1940.

[3]) The approximate constancy of this ratio of isotopes is in many cases an essential condition for the usefulness of the tracer method; all the quantitative investigations according to this method are based upon the fact that the spatial distribution of the one isotope ascertained at the end of a process is a faithful picture of the whole preparation or of the whole amount of that substance which was made to take part in the process. We shall return to this later.

in Norway, D_2O is obtained on a large scale as a by-product in the preparation of the hydrogen needed for the synthesis of ammonia; that process is carried out electrically in that country, where electrical energy is very cheap. The oxygen isotope $_8O^{18}$ is prepared by fractional distillation of water, making use of the fact that H_2O^{18} is less votatile than H_2O^{16}. The nitrogen isotope $_7N^{15}$ is obtained by a chemical process repeated many times.

The synthesis of tagged compounds

In many cases one works directly with tagged elements. Very often, however, experiments are made with compounds of those elements, which means that one actually needs "tagged compounds". How are such compounds obtained?

It is usually a question of compounds which in their untagged form could be purchased by the investigator from the chemical industry. If this is not the case, as starting materials for the synthesis he could at least choose compounds obtainable commercially which are as close as possible to the desired product and which can be converted into the latter by means of simple reactions. The tagged compounds, however, will usually have to be prepared by the investigator himself, and this will sometimes be made more difficult by the fact that he cannot start with a closely related compound but will be compelled to begin with the tagged element as such, or perhaps from the compound in which the isotope is formed in its preparation. In the case of some radioactive isotopes, such as C^{11}, there is the further difficulty of the limited lifetime, which makes it necessary to carry out the synthesis very rapidly. If, owing to such difficulties, the investigator has no success or the normal methods of synthesis are too laborious, the solution of the problem may perhaps be found in the excellent method of biological synthesis. The more or less complicated organic substances needed for many physiological investigations often occur as products of the metabolism of animals or of certain micro-organisms. If, for example, glycerine is administered to a culture of B pentosaceum (a kind of bacteria), together with carbon dioxide with radioactive carbon, within 30 minutes 80% of that carbon is obtained in the form of proprionic acid and succinic acid. With other micro-organisms other compounds can be made.

A method which can be used in many cases is to allow animals to drink "heavy water" for a time, or to water plants with heavy water. After some time numerous compounds can be isolated from the tissues in which a part of the normal hydrogen occurring in them has been replaced by heavy hydrogen.

A similar case of biological synthesis may be seen in the preparation of lecithin (a group of fats containing phosphorus) which is tagged by means of radioactive phosphorus. Interesting experiments have been performed with this substance, which has been recently much advertised as a "nerve tonic". When radioactive lecithin was fed to a rat it was found that after some time the blood of the rat became radioactive; but the radioactivity could not be localized in the lecithin which is always present in the blood plasma, but in the inorganic phosphate which is also present in the plasma. This means that the lecithin is not taken up as such through the wall of the intestines to reach the blood and organs in that way, but that the lecithin is split up in the intestines into its components. Thus, instead of taking lecithin, the same result can be obtained by consuming a mixture of fats and inorganic phosphates in the correct proportions. The body itself then builds up the lecithin. Use was made of this very fact in order to obtain the radioactive lecithin required for these tests; radioactive phosphate is injected into a hen and several days later radioactive lecithin can be isolated from the yolks of the eggs laid by the hen.

The analysis of the mixtures of isotopes

In the case of radioactive isotopes the analysis is performed by measuring the radioactivity. Since practically all the radioactive isotopes considered emit negative or positive electrons, the measurement can be performed with a so-called electron counter [4]). This instrument contains a gas-discharge tube made conductive for a very short time by an electron entering it. Thereby a current impulse is excited in the electric circuit in which the tube is connected. These impulses are amplified and counted by a counting apparatus. The number of impulses per second is a direct measure of the radioactivity of the substance placed in the vicinity of the counting tube.

It is clear that the value of the activity thus measured will still depend very much on the geometrical configuration, *i.e.* on the shape and size of the prepared object and on the relative distance and position of object and counter. It is always sufficient,

[4]) See for example A. Bouwers and F. A. Heyn, A Simple Instrument for Counting Electrons, Philips Techn. Rev. **6**, 74, 1941. In addition to electron counters ionization chambers with electroscopes are often used for measuring radioactivity. These instruments are simpler but less sensitive. The following discussion is equally valid for all measuring arrangements, so that we need only speak about the electron counter.

however, to perform relative measurements: it is a question of comparing the radioactivity of two substances or of one substance at different times. For this it is only necessary to provide for satisfactory reproducibility. In the most common case where one is concerned with preparations in powder form, it is customary to spread out the powders in a certain thickness, in a dish of certain dimensions and then to place the dish in a well defined position with respect to the counter (*fig. 2*).

duced by combustion, so that the activity of each part is concentrated in a very small volume. The activity of these different preparations can be compared by placing the same quantity in each case under the counter. For comparison with the original (very concentrated) preparation of the nutritive substance, the latter is made up, if necessary, to the required amount by the addition of inactive ash in order to obtain the same configuration for the measurement.

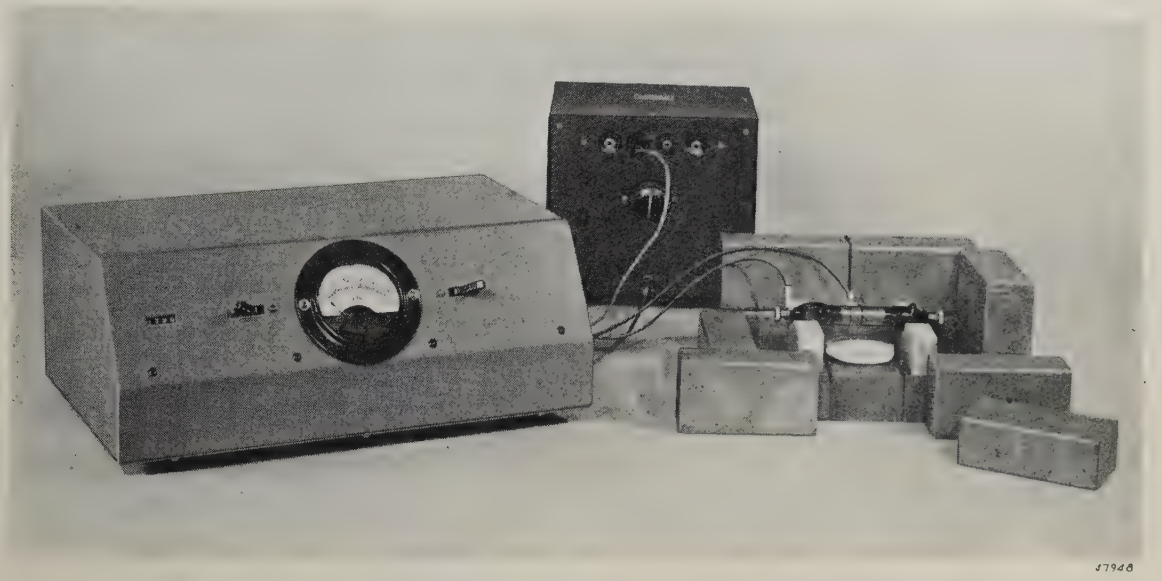


Fig. 2. Arrangement for the measurement of radio-activity with an electron counter. Under the counter tube is a dish in which the substance to be investigated is spread out (right). The two poles of the counter tube are connected with the counting apparatus seen on the left of the photo. In the middle background may be seen the feeding apparatus. Even in the absence of a radioactive substance the counter records a certain number of impulses per minute which are caused by cosmic rays and by the natural radioactivity of the surroundings. In order to diminish this so-called zero effect the counter tube with the substance is surrounded on all sides by lead blocks, some of which can be seen in the photo.

If it is only a question of localizing the radioactive substance, as in the case of the scraper stuck fast in an oil pipe line — mentioned as an example in the preceding article [1]) — or as in experiments for the assimilation of phosphorus or potassium out of the soil in different parts of a plant, the pipe line or the plant can simply be gone over with the counter. More precise localization is possible by the method of the auto-radiogram also previously mentioned. The assimilation of zinc in the fruits of a tomato plant, for example, has been studied by cutting through a fruit and laying the cut surface on a photographic plate. The spots where the fruit contained radioactive zinc are indicated on the developed plate by a blackening. The tomato was found to assimilate the zinc mainly in the seeds (*fig. 3*).

For more quantitative work in experiments like the above the different parts of the plant are re-

One complication generally occurring in the analysis of radioactive preparations is the natural decrease in the activity of a substance with time. When it is desired to calculate from the measurement of activity the fraction of the total amount of radioactive substance which has entered into the preparation to be analyzed, a factor depending upon the time should enter into the calculation. This, however, is avoided in a very simple way when the activities measured are always expressed as a percentage of the activity of a portion of the original preparation which has not been used in the experiment. Since the activity of the latter decreases at the same rate, the time factor is automatically eliminated.

The sensitivity that can be obtained with the electron counter is extremely high. With ordinary counters the presence of an amount of 10^5 radioactive atoms can be demonstrated, *i.e.* 10^{-18} to

10^{-19} gram molecules. The limit is set by the statistical fluctuations of the number of atoms disintegrating per unit of time and the fluctuations of the zero effect (see the text under fig. 2).

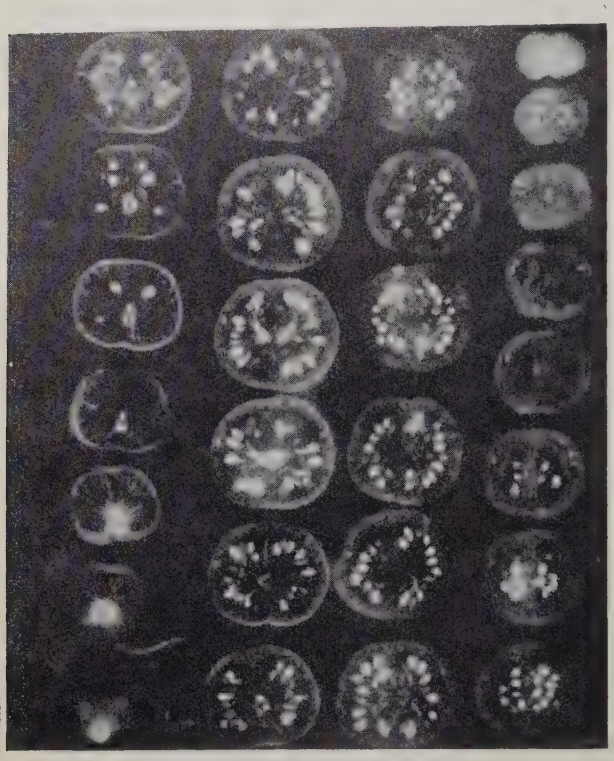
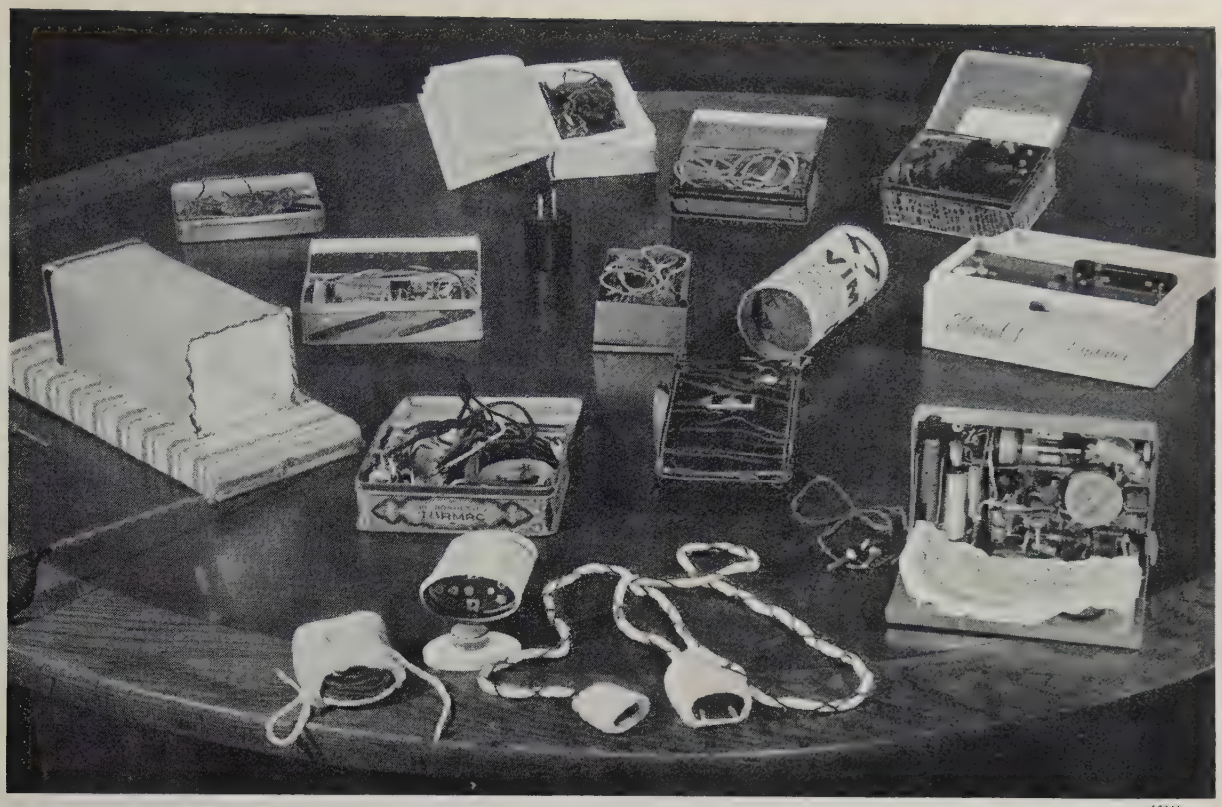


Fig. 3. Auto-radiogram of a cut surface of a tomato. The plant had been nourished for some time with radioactive zinc, which for the greater part was accumulated in the seeds.

With the preparations used in the tracer method the concentration of the radioactive isotope is usually not more than $1 : 10^{10}$ (see the article referred to in footnote [1]); it is then possible to demonstrate the presence of 10^{-9} gram molecules of the preparation. With particularly sensitive counters and very concentrated preparations it has been possible to demonstrate the presence of 2×10^{-12} gram molecules of copper, *i.e.* an amount of 10^{-10} gram.

According to the attainable reproducibility of the geometrical configuration, the accuracy of the radioactivity measurements with the counter varies between about 0.1% and 5%. In the neighbourhood of the limit of sensitivity the accuracy naturally decreases. An accuracy of 10% means that for example in exchange experiments it is only possible to determine the amounts of the isotopes exchanged with a possible error of 10%. Sometimes, however, this is sufficient for the purpose in view. The difference in mass of two isotopes is also accompanied by a slight difference in diffusion and reaction velocity (see above under the preparation of stable isotopes), so that the ratio of isotopes in a physical or chemical exchange process does not always remain entirely constant, but may change slightly, sometimes even by several percent.

The analysis of mixtures of stable isotopes is a problem by itself, which we shall only touch upon lightly here, because these isotopes are of less importance for practical applications. The oldest and in modern form still much used method of analysis is the "mass spectrograph" method of Aston. In a gas-discharge tube ions of the mixture are formed and then deflected to a greater or lesser degree by an electrical and magnetic field according to their mass. This method has been developed to a high degree of accuracy, namely to errors of only 0.5% in the ratio of isotopes to be determined. A much easier method of analysis is that based on the measurement of density, which is used for example for deuterium, the heavy isotope of hydrogen, in the form of "heavy water" (D_2O). Since it is possible without too great difficulties to measure the density of water to an accuracy of 10^{-6} — in this connection it is also important that water can so easily be purified by distillation — it is possible to determine D_2O quantitatively in extremely great dilution in H_2O. If the deuterium forms part of other hydrogen compounds the same method of analysis can be applied by first burning the compound and concentrating attention on the water formed in the combustion. Sometimes only very small quantities are then available for the measurement of density, but this is no objection since there are methods of measurement in which 1/100 or even 1/1000 cm^3 of water is sufficient.



48111

SECRET PRODUCTION OF RADIO RECEIVERS IN OCCUPIED TERRITORY

by "ONE OUT OF MANY." 621.396.181.4"36"

During the German occupation the Dutch were strictly forbidden to listen to foreign broadcasts, but as long as they were left in possession of their wireless sets, few obeyed the orders. When, however, in 1943 the German authorities enforced their commands by confiscating wireless sets on a large scale, there arose a need everywhere for pocket-size receivers, which could either be hidden away in a small corner or camouflaged in an article of daily use. It was, of course, primarily the task of the Philips concern in Eindhoven to satisfy this demand, and though a normal and organized production was out of the question, personal initiative was not lacking. Soon a lively activity sprang up in small workshops and private rooms and the amateur set-making of earlier days underwent a temporary but intensive revival.

Official production figures are lacking and no one knows exactly what was achieved, but it is estimated that several thousands and perhaps tens of thousands of these pocket-size receivers were produced in Eindhoven. Though we can only draw on our own experience and on that of a small circle of friends we hope that the following notes will nevertheless give a fair idea of the difficulties and anxious moments of all who partook in this more or less risky venture.

Preferably "button" or "acorn" valves were used to keep the set as small as possible, and the so-called reflex circuit was very popular owing to the high sensitivity to be obtained from it with only two of these miniature valves. In this circuit the first tube serves both as high frequency and low frequency amplifier, while the second tube is used as grid detector with feedback; the low frequency signal from the anode of this last valve is impressed on the grid of the first valve. In this circuiting system the H.F. and L.F. parts must be carefully separated by filters to avoid self-oscillation.

Fig. 1 illustrates an example of such a circuit, which when using a headphone would give a satisfactory reception on an aerial of two or three meters stretched out in the room.

Since there was no space for a normal tuning condenser, so-called trimmers with a capacity varying from 3 to 30 $\mu\mu$F were used (*fig. 2*). With

well chosen coils the 30, 40 and 50-meter bands could then be covered, which left an ample choice of stations to be tuned into.

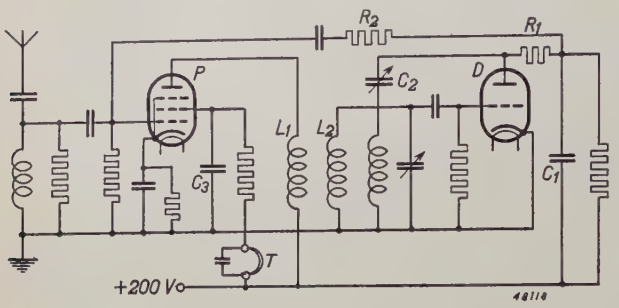


Fig. 1. A reflex circuit frequently used. The pentode P serves as h.f. and as l.f. amplifier; its anode is inductively coupled with the grid of the triode D which is employed as grid detector. From the anode of D the low frequency signal is fed back into the grid of P *via* a low-pass filter comprising the elements R_1, C_1 and R_2. The headphone is taken up in the lead to the screen grid of P, the condensor C_3 being inserted to avoid h.f. signals in the telephone leads.
The triode D has also been fitted with a normal h.f. feed-back which is regulated by the condenser C_2. A weak inductive coupling between the coils L_1 and L_2 precludes the transmission of l.f. signals from P to D.

Another circuit frequently adopted is illustrated in *fig. 3*. In this circuit the triode-hexode UCH 21 is employed, the triode being used as grid detector and the hexode as L.F. amplifier. This combination is made possible by the fact that the grid of the triode is not internally connected with one of the grids of the hexode.

Despite the wide choice of stations referred to above, listening in was often far from a pleasure owing to the many jamming stations set up by the Germans especially in the short wave range. Combined with the unavoidable fading these made reception difficult and frequently impossible, and in the course of time this induced many set-makers to change over to longer wavelengths. The smaller range of wavelengths then to be covered was compensated by a more reliable reception.

On the medium wave the British "Home Service" in particular was free from jamming for a long time, and on the 1500 meters the news could still be distinguished with sufficient clarity from a disturbing background. When, however, the station usually tuned into happened to change its wave length it became necessary to alter the coils and this was generally no easy matter with such minute home-made sets.

The circuits illustrated by figures 1 and 2 were specially designed for sets of very small dimensions. Those, however, who had given up their wireless to the German authorities, according to orders, and consequently did not live in constant fear of their house being searched, could take the risk of somewhat larger dimensions. In a collection recently exhibited to members of the B.B.C. staff visiting the Netherlands we observed a couple of receivers mounted on a sturdy brass chassis and fitted with interchangeable coils and two normal valves, the EF 6 and EBC 3, which reproduced various English programmes on a loudspeaker with an aerial of only one meter. We know that 15 of these

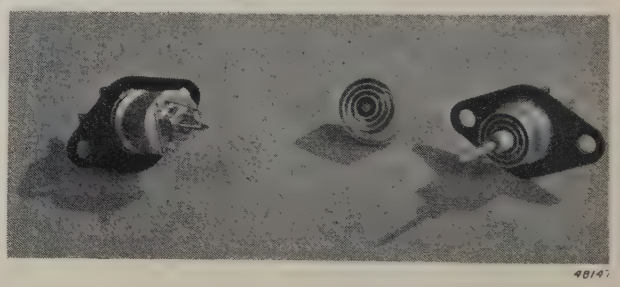


Fig. 2. A trimmer, ordinarily used for the correction of capacities in tuned circuits, is now employed as tuning condenser. The construction shown above is that of a sliding condenser, one electrode being moved with respect to the other by a screw movement; the capacity ranges from 3 to 30 $\mu\mu$F.

sets were made by one person single handed, some of which were supplied to the underground movement. And this is by no means to be considered as a record.

Practically all pocket-size receivers were made for direct use on the 220 volt AC mains, the anode

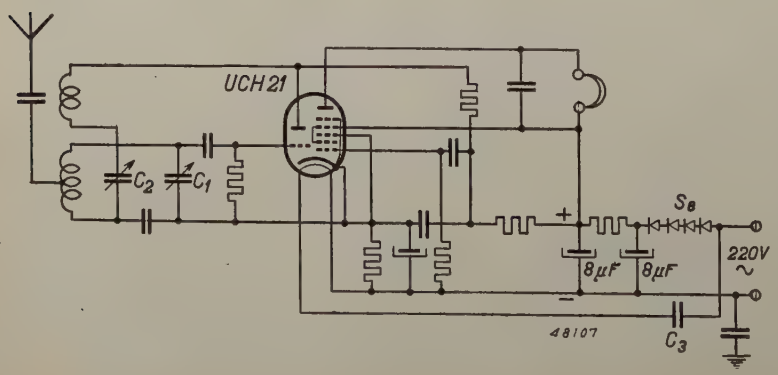


Fig. 3. Another circuit of frequent application, combining grid detection and l.f. amplification in a single valve (UCH 21). A normal feed-back can be adjusted by means of C_2. To avoid the use of a transformer the heater is fed from the mains *via* the condenser C_3. The anode voltage is also taken from the mains after rectification by the selenium rectifier *Se*.

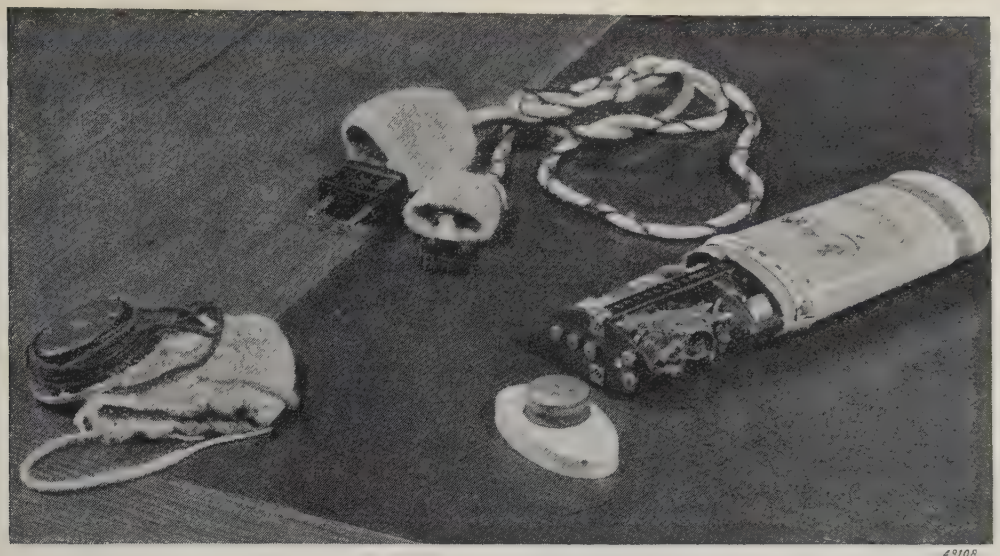


Fig. 4. A wireless set assembled in baby's powder duster. The connecting flex camouflaged by knitted covering was used as girdle on mother's dressing gown. The telephone covered in similar fashion was hung in the cradle as baby's rattle.

voltage being obtained by rectifying the mains voltage either with an "acorn" valve or with a selenium rectifier which could supply a maximum current of 10 mA. Smoothing was effected with an electrolytic condenser specially developed for this purpose comprising 2×8 μF in a cylindrical space 18 mm in diameter and 50 mm long.

The cathode heating current was in many cases furnished by a small transformer. Since, however, these were not always easy to come by, the heating current was in other cases taken directly from the mains, an incandescent lamp or better still a condenser being used as series impedance. A valve like the UCH 21, being designed for a small heater current and a high heater voltage, was specially suitable in this respect. Appropriate condensers were again constructed and manufactured on a fairly large scale solely for this purpose.

It should, however, be noted that the use of a transformer had one definite advantage in that it

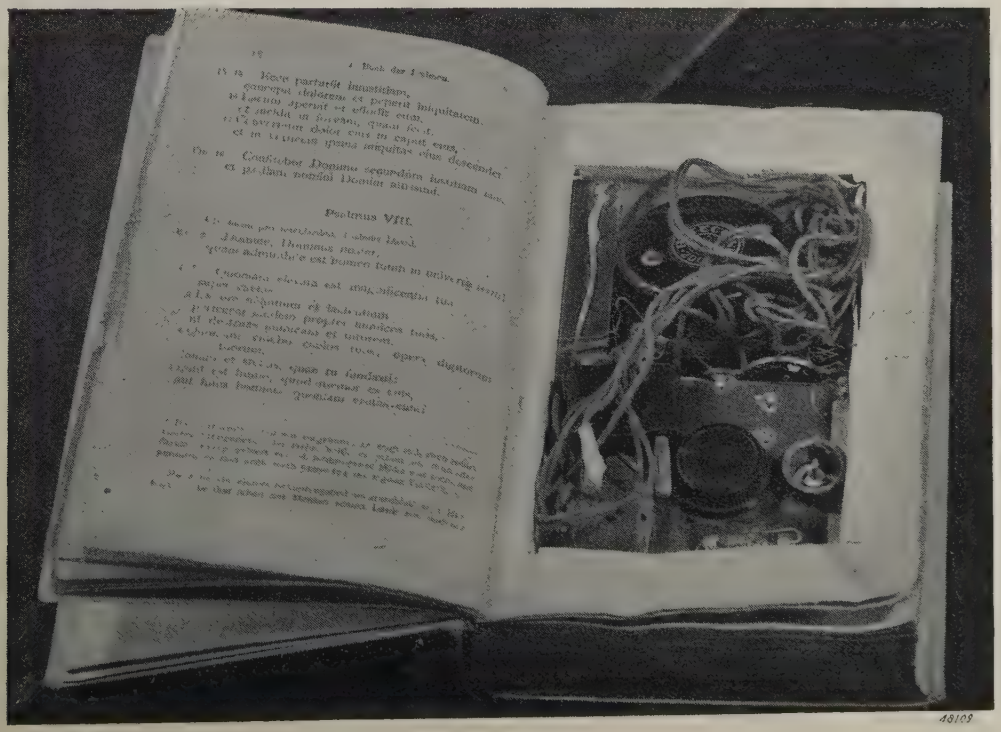


Fig. 5. A wireless set concealed in the cut-out pages of a book.

could also be fed from a bicycle dynamo, connected to the low voltage side. In this case the dynamo furnishes the 6 volts required for feeding the filaments and this low tension is stepped up by the transformer to the 220 volts needed for the anode

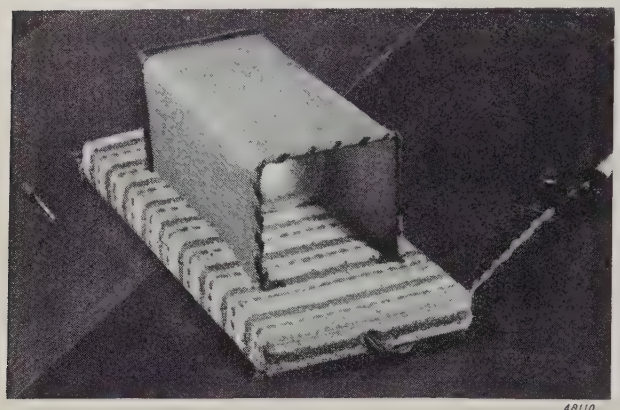


Fig. 6. The base of this lamp contains a superheterodyne receiver with 5 valves and a telephone serving as loud speaker. In addition to its obvious use as reading light the wireless could be put in operation by plugging a nail in an almost invisible hole. The spoke of a bicycle wheel serving as aerial could be slipped right back inside when not in use.

voltage. Turning the bicycle upside down and mounting the dynamo on the rear wheel fork, the set could be operated by driving the pedals by hand. In this way many have been able to get the news even when for some reason or other the supply of electricity broke down.

To collect the necessary components was not always an easy matter and frequently the construction of the set will have been prescribed by whatever materials happened to be available. Anyone having a stock of special parts under his custody would sooner or later be obliged to keep it safely under lock and key to prevent it from dwindling in a mysterious manner.

In finding a suitable camouflage individual ingenuity was given free rein. At an exhibition of the resistance movement we noted for instance a wireless receiver assembled inside the headlight of a bicycle lamp; as the glass was covered with the prescribed black-out screen nothing could be seen from outside. Other devices are illustrated in the picture heading this article, while in *figs. 4, 5 and 6* some original "solutions" are shown in greater detail.

The most urgent need for reliable news was, of course, felt by those imprisoned in German camps. In the spring of 1944 a Dutch prisoner of war in Neu Brandenburg succeeded in sending a request to Eindhoven for certain specified parts to be sent to him concealed in a food parcel. So much experience had by that time been gained that it could be decided to send him a complete set instead, camouflaged in a tin of vegetables. In this case the circuit of fig. 3 was adopted and after the set had been assembled the weight of the tin was made up to its proper value, care being taken that the centre of gravity coincided with its normal position. Furthermore the inside of the tin was provided with a special lining so that it would produce the correct sound on being tapped. When eventually a message came through that the "football boots (this was the prearranged code-word) had arrived in safety and were an excellent fit" our joy was great, as will easily be understood.

Needless to say this war-time venture had its dangerous side, and unfortunately some who took part forfeited their lives; in spite of that, however, the work went on, outwitting the enemy with cunning enterprise.

ELECTRICAL DETECTION OF TRACES OF POISONOUS GASES IN THE ATMOSPHERE

by J. BOEKE.

613.632.4 : 544.4

A catalyst of finely divided platinum which catalyzes an oxidation reaction is found to be hindered in its function by many gases which are also poisonous to the human body. Based on this fact, an instrument has been designed which can detect the occurrence of poisonous gases, for instance carbon monoxide in mines, garages or boiler-houses, hydro-cyanic acid gas in galvanizing shops, etc. The instrument contains a platinum catalyst on a platinum wire serving as one of the resistance branches in a Wheatstone bridge and along which a mixture of methylalcohol and atmospheric air is conducted. The heat developed by the oxidation of the alcohol keeps the platinum wire at a temperature of 120-150 °C, at which the bridge is brought into equilibrium. When the catalyst is poisoned the oxidation reaction ceases, the wire cools off and, due to its change in resistance, the bridge is thrown out of equilibrium, thereby setting an alarm arrangement into action. A simple portable model of the instrument is described.

Suffocation and poisoning

The energy for the vital processes in man and animal is furnished by "combustion" of the organic compounds in the food: an oxidation of the hydrogen and the carbon in these compounds by the oxygen in the air. For this oxidation it is necessary that oxygen can be inhaled from the surrounding air through the respiratory organs, eventually being carried by the blood to the various organs involved in the oxidation processes.

If, therefore, the air inhaled by an animal were replaced by a gas containing little or no oxygen, for instance only nitrogen or laughing gas (N_2O), the animal would die for lack of oxygen: it would suffocate, or some of its vital functions would cease; it would become unconscious.

Elihu Thomson showed in 1873, by experiments on the Thomson's family cat, that in the case just described the term suffocation must be used in contrast to the poisoning which occurs, for instance, due to the presence of carbon monoxide in the atmosphere. In the case of suffocation (at least if life is not already extinct), upon oxygen being administered, the organism quickly resumes its normal functions, whereas in the case of the poisoning mentioned the after-effect lasts longer and — in spite of renewed supply of oxygen — death may still set in after some time. In the case of such poisoning as this it is not the external supply of oxygen that ceases, but rather the transfer of the oxygen from the lungs to the compounds to be oxidized. The agents serving for this transfer are the haemoglobin (red pigment) in the blood — which together with oxygen forms a chemical compound, oxyhaemoglobine, that can easily be decomposed again and carries it in that form to the organs — and a number of specific substances called ferments or enzymes. The haemoglobin combines with carbon monoxide more readily than with oxygen, so that where carbon monoxide is present the oxygen-blood compound can no longer be formed and the supply of oxygen through the body stops. In a similar way the ferments in the body can be rendered inactive by certain gases in the air inhaled. Their function as carriers also ceases; they are "poisoned". Since the ferments are present only in very small amounts it is understandable that even only traces of a poisonous gas may be sufficient to block a special ferment function essential in the chain of metabolism of the organism, and thereby cause the entire organism to suffer a "death by starvation". Only 0.03 vol. % of the well-known prussic acid (hydrogen cyanide) in the atmosphere inhaled for a few moments is sufficient to kill an adult. It is less well known that hydrogen sulphide (given off by rotten eggs) is just as poisonous and in fact, because its action is cumulative, even more dangerous. Due to its extremely strong odour, however, we are warned in time and can escape from the poisonous atmosphere.

Unfortunately in many other cases of poisonous gases the nose does not function as an alarm instrument. Our sense of smell very often fails to react to gases which ordinarily do not occur free in nature but which are usually only formed by human agency: carbon monoxide is quite odourless for one, hydrogen cyanide for many. Even with respect to hydrogen sulphide, which does often occur in nature, the nose is not always reliable: after a relatively short time the sense of smell becomes so used to it that the concentration could be increased until a fatal dose is unsuspectingly inhaled.

In practice there are many cases where there is a

risk of coming into contact with these or other poisonous gases. For example, in garages or boiler-houses, where carbon monoxide may be formed owing to incomplete combustion; in mines where carbon monoxide may also occur; in painters' workshops where carbon monoxide is liberated in the drying of paint containing linseed oil; in the fumigation of ships' holds and of houses with hydrogen cyanide preparations, where traces may linger after the treatment; in galvanizing shops where hydrogen cyanide may be developed owing to incorrect handling of cyanide-containing baths, etc. In such cases it may often be desirable to have at one's disposal an objective alarm apparatus sufficiently sensitive to react even to traces of invisible, more or less odourless poisonous gases and vapours.

At first glance this might seem impossible of achievement. How can an instrument distinguish whether a gas is poisonous to human beings or not? The instrument shoud not, of course, react to non-poisonous gases.

A solution of this problem is made possible by the fact that the phenomenon of poisoning in the human body shows a far-reaching parallelism with the phenomenon of the poisoning of a platinum catalyst. An alarm arrangement based upon this will be described here [1]). For a good understanding of its working, however, it is necessary first to say something about the functioning of a catalyst.

Catalysts

Oxidation by transfer in addition to ordinary oxidation is also met with outside the living organism.

Hydrogen is an easily oxidizable gas which on being burnt gives off heat and forms water. At room temperature, however, no combustion takes place, due to the fact that between the two molecules of hydrogen and oxygen approaching each other there is an energy threshold preventing them from combining at room temperature. If the temperature is increased by the flame of a match the energy of motion of a number of molecules is increased sufficiently to cause them to pass over the energy threshold; the heat of combustion then liberated keeps the reaction going.

However, there is yet another possibility of bringing about the oxidation of hydrogen (and of compounds containing hydrogen), namely by means of a preparation of finely divided platinum, so-called platinum black, upon which both hydrogen and oxygen are absorbed. This phenomenon, which was discovered as early as 1823 by Döbereiner, may be explained as follows. In the adsorbed state the molecular structure of the gases is somewhat loosened by the fields of force of the platinum atoms, which means a lowering of the above mentioned energy threshold, so that the latter can already be exceeded at room temperature and reaction takes place. The platinum itself does not take part in the reaction, it only catalyses it. This may be expressed by saying that the oxygen, which in a free state does not react with free hydrogen, is carried over to the hydrogen by the platinum, acting in much the same way as the organic carriers, the ferments, enzymes, etc., which may also be considered as catalysts. Since the platinum does not take part in the reaction and thus is not consumed, and the adsorbed molecules after their mutual reaction leave the surface to make room for the adsorption of new molecules, a small amount of platinum can catalyse the reaction of any desired large amount of gas, as is the case with ferments.

Upon platinum black being heated to a high temperature its catalytic properties are lost. This can be explained by the assumption that the catalysis does not take place uniformly over the whole platinum surface, but at special "active" spots, for instance the edges and corners of the minute crystals of platinum on the surface [2]). This assumption agrees entirely with the explanation of catalysis by adsorption, since the platinum atoms at such spots are surrounded by fewer neighbouring atoms than in the plane face, and therefore have left over, as it were, forces of attraction to attract foreign molecules (see *fig. 1*). At a high temperature

Fig. 1. Diagram of the surface of a catalyst with the fields of attraction shaded (*a*). Edges and corners of crystals in the surface form active spots where foreign molecules are preferentially adsorbed (*b*).

recrystallization takes place, the many small crystals combining to form a few large ones. The number of edges and corners is thereby considerably reduced, until the number of active spots is too small to make any catalytic action perceptible.

Döbereiner, who also knew of this phenomenon, noted that traces of hydrogen sulphide, hydrogen cyanide and carbon monoxide also made platinum black ineffective; the catalyst was "poisoned".

[1]) Netherlands patent no. 54716, applied for 23 Dec. 1939.

[2]) See the survey: E. K. Rideal and H. S. Taylor, Catalysis in Theory and Practice, Mac Millan & Co. London 1926.

The explanation is that the poisonous substances are very easily adsorbed on the platinum, their molecules occupying the active spots and staying there, thus blocking the catalytic function. Since the adsorption is a surface effect, it is also clear that a very small volume of the poison can make the catalyst ineffective, since every active spot need only be occupied by one or a few poison molecules.

For reactions in which different gases are involved, different catalysts are generally used, and the poisons also seem to be more or less specific: Hydrogen sulphide, a strong poison for platinum, is not poisonous for nickel sulphide catalysts, which are used technically in the preparation of synthetic petrol (Am. gasoline). Oxygen, a normal component in reactions accelerated by platinum, is a strong poison for the iron catalysts in the synthesis of ammonia, etc.

There is, however, a remarkably pronounced parallel between the poisoning properties of all kinds of substances for platinum and for the organic catalysts which play a part in the body. Let us compare, for instance, platinum with catalase, an enzyme occurring in most body liquids such as blood and milk. Both substances may function as accelerators of the reaction in the decomposition of hydrogen peroxide into water and oxygen. It has been studied how this catalytic function is hindered by certain "poisons"; in the table below [3]) the number of gram molecules of a series of poisons is given which is necessary to reduce to one half the velocity of the reaction mentioned for a given platinum or catalase preparation. It may be seen that platinum is poisoned by a whole series of substances which are also poisonous to catalase. Carbon monoxide is, it is true, not poisonous for catalase, but it is important that platinum is also sensitive to it, since, as has already been mentioned, carbon monoxide is poisonous for haemoglobin.

It is impossible to state exactly upon what the parallelism between the poisonousness for platinum and that for the organic catalysts is based. It may be noted, however, that they both have the property of catalysing reactions at low temperature where oxygen is involved. The inorganic catalysts which catalyse the same reactions at higher temperatures, such as certain oxides, are insensitive to many of the poisons listed, probably because of the fact that at higher temperature the poison molecules themselves begin to react and therefore do not continue to block the active centres.

[3]) Borrowed from an article by W. D. Bancroft in the book referred to in footnote [2]).

Poison	Concentration of poison (in gram molecules) for a certain platinum preparation	for a certain preparation of blood catalase
Sublimate, $HgCl_2$	1/2 000 000	1/2 000 000
Hydrogen sulphide, H_2S .	1/300 000	1/1 000 000
Hydrocyanic acid, HCN .	1/20 000 000	1/1 000 000
Mercuric Bromide, $HgBr_2$	—	1/300 000
Mercuric Cyanide, $Hg(CN)_2$	1/200 000	1/300 000
Iodine solution, I_2 in KI .	1/50 000 000	1/50 000
Hydroxylamine hydrochloride, $NH_2OH.HCl$.	1/25 000	1/80 000
Phenyl hydrazine, $C_6H_5 NHNH_2$	—	1/20 000
Aniline, $C_6H_5NH_2$	1/50 000	1/400
Arsenic trioxide, As_2O_3 . .	1/50	non-poisonous to 1/2000
Carbon monoxide, CO . .	very poisonous	non-poisonous
Hydrochloric acid, HCl .	1/3000	1/100 000

Principle of the alarm apparatus

The fundamental principle of the alarm apparatus, which is based upon the above parallelism, is the following. An oxidation reaction is caused to take place on a platinum catalyst. Heat is thereby liberated. The oxygen for the oxidation is furnished by suction of atmospheric air. If the latter contains a poison (in the form of gas, vapour or mist) the catalyst loses its effect, the reaction stops and thus also the heat development. The check on poisonous substances in the air is thereby reduced to a check on the heat development.

As oxidation reaction we have chosen the oxidation of methylalcohol. A mixture of methyl alcohol vapour and air flows along a layer of platinum black deposited on a platinum wire with an intermediate layer of aluminium oxide. This wire forms one of the arms of a Wheatstone bridge (*fig. 2*). By the heat of oxidation released by

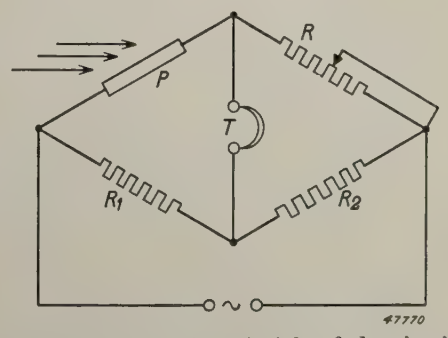


Fig. 2. Diagram showing the principle of the circuit of the alarm apparatus. The platinum wire P bears the catalyst (a layer of platinum black) along which a mixture of methylalcohol and atmospheric air is blown. R variable, $R_{1,2}$ fixed bridge resistance, T headphone.

the catalyst the temperature of the platinum wire and thus its resistance is increased. The bridge is brought into equilibrium at the temperature ultimately attained by the wire after the oxidation has proceeded for several moments (120-150 °C).

The presence of a poison will immediately cause the temperature of the catalyst and its carrier to fall, the equilibrium of the bridge is disturbed and a buzz is heard in the head phones shown in the diagram. Of course it is also possible to cause the disturbance of the equilibrium of the bridge to actuate a series of sirens or signal lamps by means of a relay.

After an alarm has been given the apparatus is prepared again for detection by "depoisoning" the catalyst. This is done by raising its temperature for a short time to about 400 °C, which evaporates the poison or causes it to disappear owing to a reaction with the mixture of methyl alcohol and air.

Activity and equilibrium temperature of the catalyst

In order that the catalyst should be hindered in its function by very small amounts of poison, it is desirable that the activity of the catalyst (*i.e.* the number of active spots) should be small compared with the amount of the reaction components passed over it per unit of time. This means that all the active spots available are continually occupied with the reaction. If this were not the case the blocking of a certain number of active spots due to occupation by poison molecules would not necessarily affect the total amount of reaction product formed per unit of time. The heat developed per unit of time would then also remain unchanged, and with certain conditions of cooling (by radiation and convection) the temperature, which is determined by the equilibrium between heat developed and heat dissipated, would not necessarily undergo any change either.

The activity of the catalyst must therefore expressly be reduced until it operates under full load. This is done by heating the catalyst wire for some time at about 700 °C. Obviously, however, one must not go beyond a certain limit, because an entirely inactive catalyst can no longer be neutralised by a poison.

This limitation can be expressed in another way. As the activity of the catalyst is reduced so the temperature at which it works must increase in order that the oxidation reaction may keep itself going. In the extreme case of zero activity, the working temperature must simply be equal to the ordinary flame temperature at which the oxidation can keep itself going without a catalyst. From this it follows that with increasing equilibrium temperature of the catalyst there can be less and less question of "poisoning": the parallelism with the sensitivity to poisons of the enzymes and ferments which act at blood temperature becomes less and less pronounced.

An equilibrium temperature of 120 to 150 °C was found to be the most suitable. In making this choice it must also be taken into account that the working temperature should not be too low, in order to ensure adequate disturbance of the bridge equilibrium when the catalyst becomes poisoned and cools down to room temperature, so that a clear indication can be obtained with a very simple and not highly sensitive apparatus.

The fact that we may count on a cooling to room temperature may be explained as follows. Upon poisoning, *i.e.* reduction in the number of active spots available, less heat will be produced by the reaction. On the other hand, as already stated, to maintain the reaction with a smaller number of active spots a higher temperature is necessary. From this it may be seen that upon a slight poisoning of a fully loaded catalyst no new equilibrium can be established; the temperature falls rapidly to room temperature.

The desired equilibrium temperature of 120-150 °C is realized by annealing the catalyst until with the given glow of the alcohol-air mixture it adjusts itself to that temperature. Of course the velocity of supply of the reaction components must then remain constant, in order that the equilibrium of the bridge shall not be disturbed already without any poisoning. The regulation of the velocity of flow is found not to be very critical, since the influence of the more rapid supply of the reaction mixture as the flow increases is approximately compensated by the deeper cooling then taking place. It is important, however, to keep the concentration of alcohol in the mixture very constant.

When the catalyst is put into action the oxidation reaction does not start spontaneously. The catalyst must first be heated artificially for a moment above the equilibrium temperature to be reached. This is done in the same way as the "depoisoning" previously mentioned.

Construction of a portable instrument

A portable alarm apparatus constructed according to the principles described is shown in *fig. 3*. In addition to great sensitivity, reliability and simplicity of operation, it was necessary — to make it portable — to aim at light weight and small energy consumption. The solution was found in the following details.

The catalyst wire is mounted in an exchangeable tube in the same way as the metal filament in an incandescent lamp.

The air to be tested is supplied by means of a diaphragm pump. This consists of a flat loudspeaker diaphragm with a cover in front of it with a blast opening. The diaphragm is vibrated by a kind of electric bell. This simple arrangement works as a pump with a capacity of about 7 liters air per minute.

Via a needle valve and a capillary methyl alcohol drips from a reservoir into a mixing chamber where the alcohol vapour is carried along by the current of air passing through. A fine wire in the capillary permits a fine regulation of the alcohol supply and a rapid cleaning of the capillary if it becomes stopped up. The mixing chamber is pre-heated by the exhausted oxidation gases; thanks to this measure the working temperature of the catalyst is little affected by temperature variations of the outside air.

In order to obtain a buzzing tone in the head-

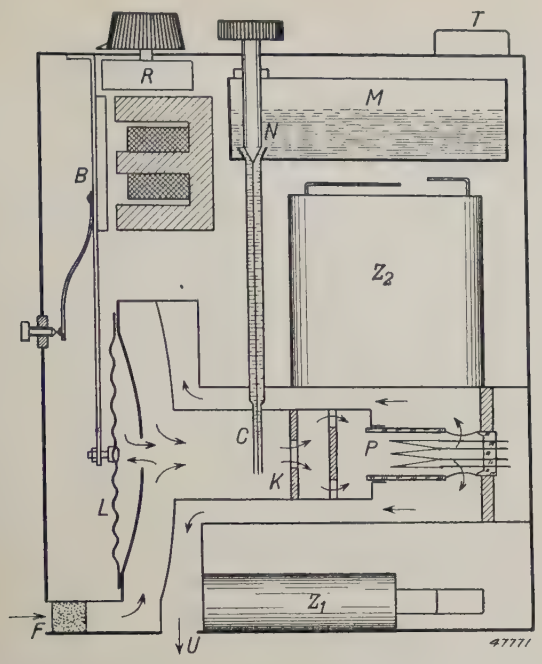


Fig. 3. Cross-section of a portable model of the alarm instrument. *P* exchangeable "lamp" with platinum catalyst, *M* reservoir with methylalcohol which, via the needle valve *N* and the capillary *C*, drips into the mixing chamber *K*. *B* electric bell which sets the loudspeaker diaphragm *L* vibrating by means of which the air entering, possibly via the filter *F*, is blown through the mixing chamber. The exhausted air leaves the apparatus at *U*. *R* variable resistance for balancing the bridge circuit. Z_1 pocket torch battery for normal operation, Z_2 battery for starting and depoisoning the catalyst, *T* connection for the headphone.

phone used as indicator the resistance bridge has to be fed with A.C. voltage or intermittent D.C. voltage. For this purpose the voltage across the winding of the electric bell system driving the pump is used.

The resistances in the arms of the bridge take up energy. Since the energy consumption must be a minimum, the winding of the bell system, in which in any case current must flow, is provided with a middle tap, and the two halves are used as fixed bridge resistances. Moreover with the help of a condenser the winding is brought into electrical resonance with the characteristic frequency of the interruptor of the bell system; see *fig. 4*. By these measures the total consumption of energy of the apparatus is reduced to 0.5 Watt, which is supplied by a pocket-torch battery.

When the variable resistance for setting the bridge in equilibrium is turned completely back, a switch connects the catalyst wire with another battery, so that it can be heated for a few seconds by 3 Watts to about 400 °C, for "starting" and "depoisoning".

In front of the openings through which the pump draws in air from the outside a filter can, if desired, be placed to block the entrance for certain gases. In this way it is possible to demonstrate the presence of certain gases selectively, for instance in garages where petrol vapour is alwasy present and where the instrument is not required to react to that but only to possible carbon monoxide; in this case, therefore, a filter must be used which absorbs petrol vapour but allows carbon monoxide to pass through. The fact that the catalyst is also poisoned by petrol vapours which are only slightly harmful for the body is one of the deviations from the above-mentioned parallelism. However, it is also possible to take advantage of this fact in order to detect petrol or other vapours in places where their presence would lead to the formation of explosive mixtures. An analogous possibility of application is based on the fact that the catalyst is sensitive to Freon and other voltatile halogenated hydrocarbons only slightly harmful for human beings.

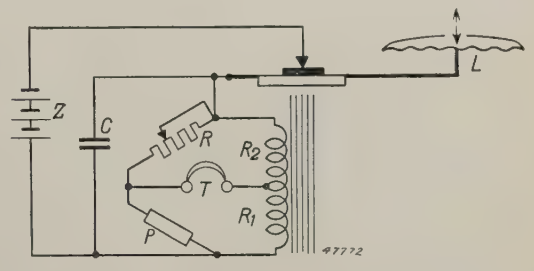


Fig. 4. More detailed diagram of the circuit of the apparatus. The two halves of the exciter coil of the bell system also serve as the fixed resistances R_1, R_2 of the bridge circuit. *Z* battery, *C* condenser, *L* diaphragm pump; other letters as in fig. 2.

These fairly expensive gases are used, for example, in refrigerating plants, and the instrument described would be very well suited for detecting leaks in such plants. It is clear, however, that such applications have nothing to do with the original purpose of the apparatus and that other specific actions or reactions of the gases in question could probably also be used for the purpose.

POTENTIAL DISTRIBUTION AT THE IGNITER OF A RELAY VALVE WITH MERCURY CATHODE

by N. WARMOLTZ.

621.316.546.032.43:621.317.329

In order to obtain a better insight into the manner in which a cathode spot is formed on mercury by means of a dielectric igniter it was desired to know, among other things, the field strength on the surface of the mercury. This article describes how this was determined with the help of a one-hundred times enlarged model of an igniter in an electrolytic tank. The presence of two dielectrics between the mercury of the cathode and the conducting part of the igniter (namely vacuum and quartz) makes it necessary to use two electrolytes in the electrolytic tank. The liquid-tight partition between these electrolytes must be conductive in the direction perpendicular to its surface, but insulating in the horizontal direction along that surface. A simple solution is given for the problem of constructing such an anisotropic wall.

When assuming that the mercury surface is smooth the results of the measurement lead to field strengths which are too low for cold emission. If, however, it is assumed, with Tonks, that the mercury surface may be drawn out to small sharp points by the electric field, then a satisfactory explanation is found for the emission observed.

A form of discharge much used in technology is the arc. Its wide-spread use is due to its power of conducting large currents with relatively low voltage loss. This is one of the factors that have made it possible to construct rectifiers of high efficiency.

Arc discharges may be divided into several kinds. For rectifiers and related applications the following are mainly used:

1) the thermal arc, where the cathode is raised to such a temperature as to bring about a thermionic emission of the order of the arc current;
2) the field arc, where electrons are liberated from the cathode by an electric field — which must be of a strength at least several times 10^7 V/cm — set-up by the positive ions formed in the arc. By way of contrast with thermionic emission one usually speaks here of cold emission.

It is this cold emission with which we are concerned in mercury-cathode rectifiers. Here the cathode is a pool of mercury and the discharge emanates from the so-called cathode spot, *i.e.* a highly luminescent spot where there is a strong concentration of electrons and ions and where the current density is estimated to average 4000 A/cm^2 (estimations vary, however, from 1700 to 18 000 A/cm^2). This spot moves about over the mercury in a most irregular manner.

Whether the mercury arc is a thermal or a field arc cannot well be decided from direct temperature measurements because of the many difficulties connected with such measurements in this case. The fact that it cannot be a thermal arc, however, is evident when it is realized that for thermionic emission of the current densities just mentioned a temperature of at least 3500 °C would be necessary; the occurrence of such temperatures must be considered out of the question because mercury very readily evaporates at a much lower temperature. From the current density and the amount of mercury that evaporates it follows, moreover, that the temperature of the cathode spot must be between 200 and 300 °C, at which temperature thermionic emission is quite negligible.

Methods of initiating a cathode spot

A cathode spot is not formed spontaneously. In order to make a mercury-cathode rectifier function, a cathode spot must be formed in some way or other. An old but still used method consists in sending an auxiliary current through the mercury cathode and an igniting electrode which reaches into the mercury and is then drawn out of it, for instance by an electromagnet or a bimetal relay. The auxiliary current is then interrupted and a cathode spot is found to have been formed. From this an arc discharge to the main anode(s) can develop.

Another method of ignition originated by Cooper-Hewitt dates from 1901. The glass wall of the vessel containing the cathode mercury is surrounded by a metal band at the level of the meniscus. When a voltage of about 10 kV is applied between the cathode and this "ignition band" — the latter must be positive — a spark occurs at the edge of the meniscus which can pass over into a

[1]) See for example M. J. Druyvesteyn and J. G. W. Mulder, The Physical Basis of Gas-filled Rectifiers with Hot Cathode, Philips Techn. Rev. 2, 122, 1937.

cathode spot [2]). This dielectric ignition, in contrast to the previously mentioned ignition by immersion with its mechanical inertia, has the advantage that the ignition can be repeated at a rapid rate and brought about at accurately determined moments. A discharge tube with dielectric ignition thus has the character of a relay valve [3]), *i.e.* a switch with practically no time lag. Moreover, thanks to the mercury cathode, it can pass very large currents without harm. These properties are of particular importance for switching devices serving to generate current impulses of variable length, interval and amplitude. The technique demands these properties, for instance, for electrical resistance welding [4]) and for stroboscopic purposes. Thus, for example, in the stroboscope [5]) already discussed in this periodical the

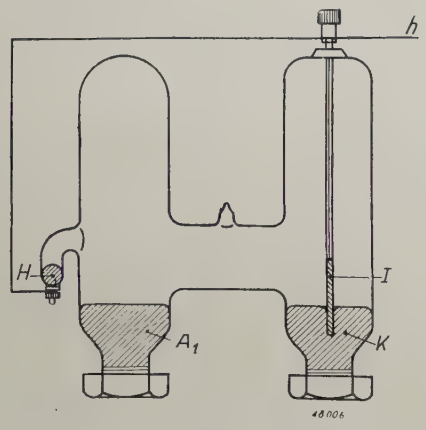


Fig. 1. Relay valve with dielectric igniter. A_1 main anode, K mercury cathode, I dielectric igniter, H auxiliary anode, h point to which a voltage impulse positive with respect to K must be applied to obtain a cathode spot.

tube sketched in *fig. 1* is used as switch for the lamp producing the light flashes. In this case current impulses of up to 2000 A are required, with a flash time of the order of 10^{-5} sec and a frequency of up to 250 flashes per sec. The ignition electrode of the tube shown in fig. 1 is an improved design of Cooper Hewitt's ignition band; it consists of a small quartz tube filled with mercury, which is shown in *fig. 2* enlarged and in a slightly different form. The action is the same as with the ignition band, the improvement consisting mainly in the

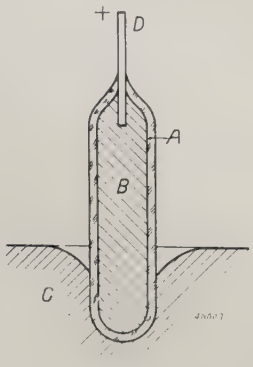


Fig. 2. Dielectric igniter. A insulating wall, B conductor, C mercury cathode, D voltage connection of the igniter.

greater durability of quartz. During the first few moments following the formation of the cathode spot an auxiliary electrode (fig. 1) connected to the same voltage source as the igniter S serves as anode for the discharge, which then develops further toward the main anode. The whole breakdown takes place within a very short time (of the order of 10^{-7} sec. calculated from the moment at which the cathode spot was formed).

When studying the properties of this igniter it was found desirable to know the distribution of potential in the space between the mercury surface and the outside wall of the igniter (respectively C and A in *fig. 3*). This is connected with the above-mentioned cold emission of electrons from

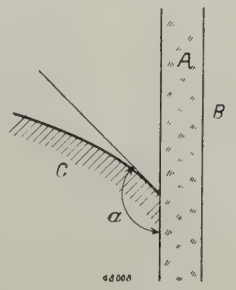


Fig. 3. Position of the mercury meniscus at the igniter. The letters A, B and C have the same significance as in fig. 2. α is the so-called limiting angle whose magnitude differs for different substances and which for the combination mercury-quartz was 136°30′.

the mercury surface. For this, of course, the electric field strength at the surface and the potential difference through which an electron passes on its path through the above-mentioned space are of great importance. These two quantities can be

---

[2]) N. Warmoltz, On the Mechanism of the Capacitative Igniter and of the Resistance Igniter in Mercury-Vapour Rectifiers, dissertation Delft 1946 (English summary).

[3]) There is also another group of relay valves, namely those with a control grid, in which the anode current can only flow when the control grid voltage exceeds a certain value. The cathode in this case may be either of mercury or it may be a hot cathode. Applications of relay valves with hot cathodes are described, for instance, in the article referred to in footnote [4]); in principe, however, relay valves with capacitive ignition can also be used there. We cannot here give a comparison of the two systems: relay valves with control grid and those with capacitative ignition.

[4]) D. M. Duinker. Relay Valves as Time Switches for spot welding, Philips Techn. Rev. **1**, 11, 1936.

[5]) S. L. de Bruin, An Apparatus for Stroboscopic Observation, Philips Techn. Rev. **8**, 25, 1946.

derived from the distribution of potential, which we have determined in the manner described below.

We did not try to calculate this distribution of potential because of the mathematical complexity of the problem. Measurements led more quickly to the goal, and in particular measurements carried out on a similarly shaped model in an electrolytic tank trough, especially since the necessary apparatus was for the main part already available [6]).

The electrolytic tank

Use is made of the fact that the distribution of potential in an electrostatic problem is identical with that in the corresponding conduction problem, that is to say it is identical with the case where the same electrodes with the same mutual potential differences are surrounded by conductors (preferably liquid, such as electrolytes) instead of dielectrics. The specific resistance of these conductors must be high compared with that of the electrodes. Both cases are governed by Laplace's equation and the same boundary conditions.

In a cross-section of the conduction model — which can often advantageously be constructed as a proportional enlargement of the original electrostatic case — the potential is measured at any desired point with the help of a moveable probing electrode. Providing the latter is small enough, it causes no appreciable disturbance of the original potential distribution.

In the practical application a tank of insulating material is used, filled with a weak electrolyte, tap water for instance. The model of the electrodes, made for instance of copper, is placed in this vessel, and voltages are applied between the electrodes (A.C. voltages of not too low frequency in order to avoid polarization at the electrodes). With the moveable probing electrode points are found where the potential has a certain value, ascertained by means of a bridge circuit. A stylus coupled with the probing electrode records the corresponding points on a sheet of paper; the geometrical position of these points is one of the desired equipotential lines.

Design of a tank with two electrolytes

In order to keep the model simple we have considered the distribution of potential in the case of plane igniters. Those actually used are in fact cylindrical, but the ignition mechanism is in principle the same in both shapes.

A possible form of a plane igniter is sketched in

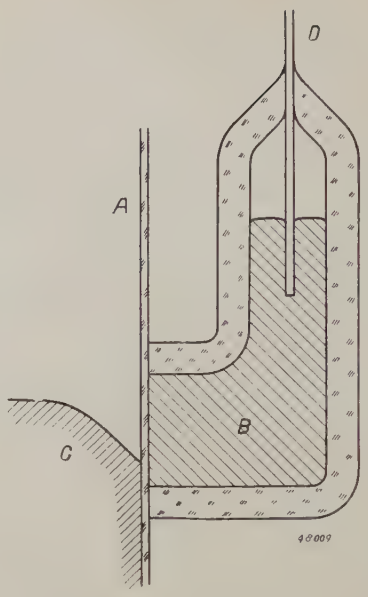


Fig. 4. Cross-section of a plane igniter. *A* quartz plate, on which the ingition occurs, *B* inner conductor (for instance mercury), *C* mercury cathode, *D* voltage connection.

fig. 4. A model of this was constructed with a hunded-fold linear enlargement for the measurements in the electrolytic tank a top view of which is given diagrammatically in *fig. 5*. The conducting part of the igniter is represented by the flat plate *B*, the mercury surface by the plate *C* bent to correspond to the meniscus, both being of copper. The space indicated by *A* corresponds to the insulating covering of the igniter. Here, however, a complication arises which did not occur in the tank measurements described in the article referred to in footnote [6]).

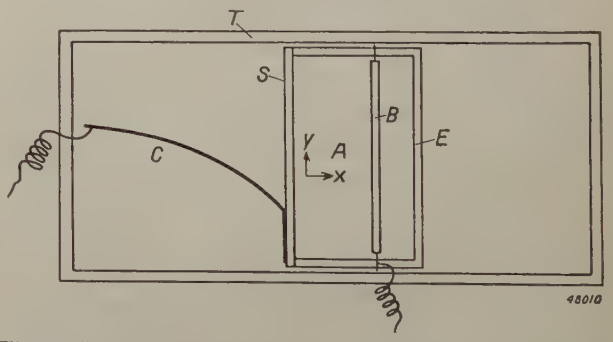


Fig. 5. View from above (diagrammatic) of an electrolytic tank (*T*) for measurement of the distribution of potential on a model of a dielectric igniter. The flat copper plate *B* and the bent copper plate *C* represent (enlarged 100 times) the analogen of the inner conductor and the mercury meniscus respectively (cf. fig. 4). The plate *B* is situated in a vessel *E* filled with a solution of copper sulphate; outside this vessel there is tap water. The space *A* corresponds to the quartz plate *A* of fig. 4. By parallel displacement of the plate the width of *A* can be *B* varied to correspond to different thickness of wall of the igniter. *S* is an anisotr copic wall conductive in the direction *X*, but not in the direction *Y* (see fig. 6). The probing electrode, which is not drawn, is moveable in the spaces *C-S* and *S-B*. The probing electrode is a copper wire of 0.1 mm diameter, projecting 1 to 2 mm out of an insulating covering.

[6]) G. Hepp, Measurements of Potential by means of the Electrolytic Tank, Philips Techn. Rev. 4, 223, 1939.

There, in the electrostatic case, there was only one dielectric (a vacuum), between the electrodes and this could therefore be represented in the tank model by one electrolyte. In our case, however, there are two dielectrics: the space between A and C (fig. 3), in which there is only mercury vapour, and the insulator A consisting of quartz or suchlike material. This means that at the corresponding places in the model different electrolytes must also be used, with conductivities in the same ratio as the dielectric constants of the insulators, *viz* 1 for vacuum (in this respect the mercury vapour present is without significance) and 4.4 for quartz. The place in the tank corresponding to the quartz (A in fig. 5) must therefore be filled with an electrolyte having a conductivity 4.4 times that of the place corresponding to the vacuum. We chose for the latter tap water and for the former a solution of copper sulphate. Since the two electrolytes must of course be kept separate, the copper sulphate is in a separate vessel.

When considering what material should be used for the wall S (fig. 5) of this vessel, one encounters the following peculiar difficulty. Metal cannot be used because conductive walls would form equipotential surfaces for which there is no analogy in the electrostatic case; this would lead to an entirely erroneous picture of the original distribution of potential. If, however, the whole vessel were made of non-conductive material, an entirely different potential distribution would again be obtained: both electrolytes would then function as coatings of a condenser, with the walls of the vessel as dielectric taking up practically the whole potential difference applied. What is desired is a construction in which the wall S (fig. 5) conducts in the direction X but not in the direction Y (its behaviour in the third direction is a matter of indifference), while the other walls must be insulating.

An obvious solution satisfying these apparently contradictory requirements is the following. A large number of metal pins are driven through an insulating plate in such a way that they project on both sides but make no contact with each other. These pins then give the desired conductivity in the X-direction, while no current can flow in directions perpendicular to that. To make such a wall, however, is rather laborious. A more practical solution has been found in the construction illustrated in *fig. 6*, consisting of a plate of insulating material surrounded by a series of spaced, U-shaped pieces of bare copper wire. The latter provide for the necessary conductivity in a direction perpendicular to the surface of the plate (X direction in fig. 5). Since the U pieces are spaced, there is no conductivity in the longitudinal (Y) direction.

It is only in the immediate vicinity of this anisotropic wall that the discontinuities of the "trelliswork" are manifested. Therefore no measurements were taken at distances less than the thickness of the wire of the spacing (both 0.5 mm).

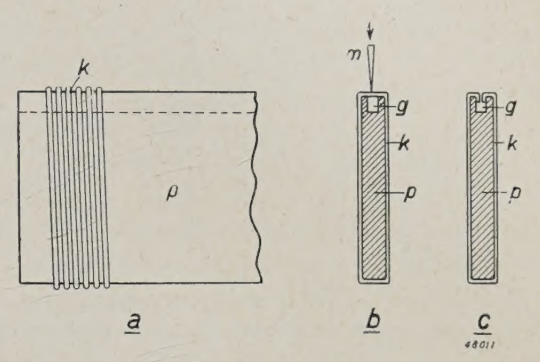


Fig. 6. Anisotropic wall separating the electrolytes (S in fig. 5). A plate p of insulating material is wound at intervals with bare copper wire k, fig. 6*a*. In the longitudinal direction there is a groove g, fig. 6*b*. With a knife m the windings are cut through above this groove and the ends thus formed are bent around the edge of the groove, fig. 6*c*. In this way a plate is formed which is surrounded by a series of spaced U-shaped pieces of bare wire.

The copper plate B (fig. 5) is moveable in the vessel containing the copper sulphate, so that measurements can be made with different widths of the space A (corresponding to the different thicknesses of quartz in the igniter).

Results of measurements

As a check, several potential distributions were first measured with tap water also in vessel E (fig. 5). It is then to be expected that the equipotential surfaces on the partitioning wall (S, fig. 5) will be continuous. *Fig. 7* shows the result of a measurement for two thicknesses of wall (0.35 and 3.0 mm, *i.e.* 3.5 and 30 cm in the model). As may be seen, by a parallel displacement over the thickness of the partitioning wall S (or rather slightly more than that thickness in order to exclude the region where the field disturbance of the individual "bars" is appreciable) the equipotential lines can be made to run almost without interruption.

Fig. 8 gives some of the actual measurements taken on the model of the igniter for three thicknesses of wall (0.35, 1.0 and 3.0 mm). The kink observed in the equipotential lines actually corresponds to the dielectric constant of the igniter wall of about 4.4.

---

[7]) This method has also been applied by R. Stachowiack, E.T.Z. **62**, 441, 1941.

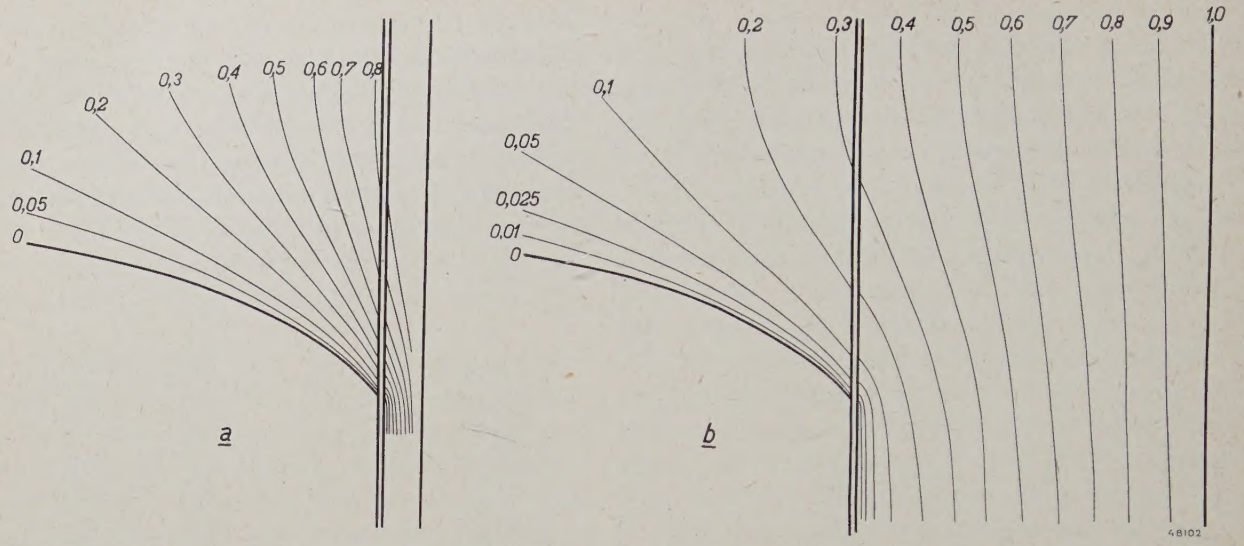


Fig. 7. Result of check measurements in which the vessel E (fig. 5) was also filled with tap water, to correspond to an imaginary igniter wall with a dielectric constant of 1 and a thickness of 0.35 mm (fig. 7*a*) and 3.0 mm respectively (fig. 7*b*). The curves are measured equipotential lines, the numbers indicating the corresponding potential (that of the electrodes C and A being set at 0 and 1 respectively). When the thickness of the separating wall S is imagined to be reduced to zero, the equipotential lines continue without interruption, as was to be expected.

The electrical lines of force undergo a change of direction at the boundary surface between two media of different dielectric constant (ε_1, ε_2), in such a way that the electrical field strength on either side of the boundary surface has the same tangential component, but that the normal components are inversely proportional to the dielectric constants. From this it follows that for the change of direction of the lines of force (*cf. fig. 9*)

$$\frac{\operatorname{tg}\beta_1}{\operatorname{tg}\beta_2} = \frac{N_1Q_1/P_1N_1}{N_2Q_2/P_2N_2} = \frac{1/P_1N_1}{1/P_2N_2} = \frac{\varepsilon_1}{\varepsilon_2}.$$

For the equipotential lines intersecting the lines of force at right angles, the following is therefore valid:

$$\frac{\operatorname{tg}\gamma_1}{\operatorname{tg}\gamma_2} = \frac{\operatorname{cotg}\beta_1}{\operatorname{cotg}\beta_2} = \frac{\varepsilon_2}{\varepsilon_1},$$

when γ_1 and γ_2 are the respective angles which the equipotential lines make with the normal to the boundary surface.

By drawing the lines of force in these figures and measuring along them the distances of the equipotential surfaces the electrical field strength can be determined. In order to find the field strength at the mercury surface several equipotential surfaces running close along the electrode C were accurately measured once more.

It was found, as is indeed well known, that the electrodes must be absolutely clean. If that is so, then the probing electrode takes up the potential of the electrode by just touching it. Otherwise there

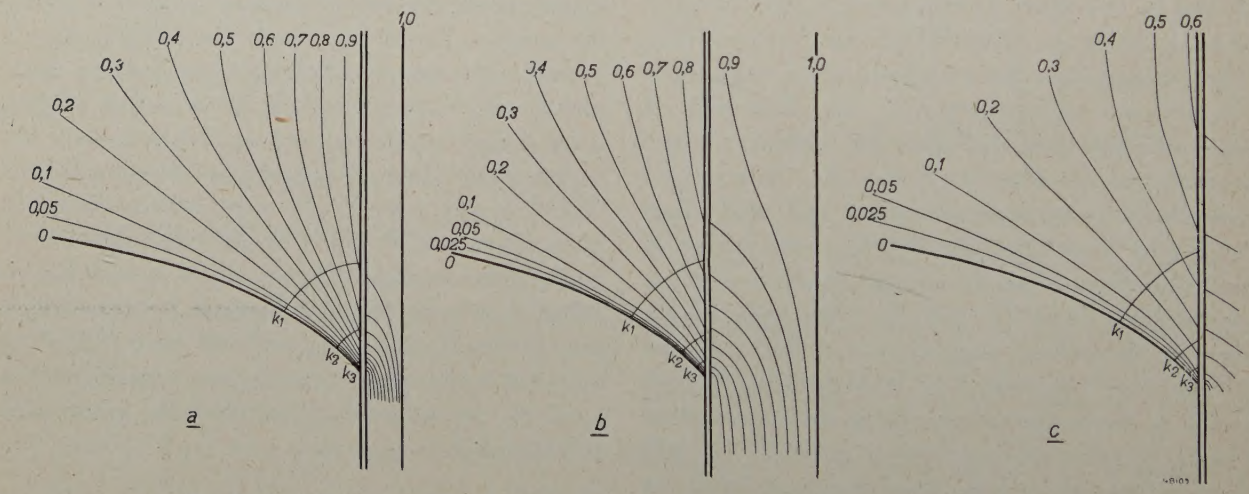


Fig. 8. Several measured distributions of potential on a plane igniter whose wall has a dielectric constant of 4.4 and a thickness (*a*) of 0.35 mm, (*b*) of 1.0 mm, and (*c*) of 3.0 mm; k_1, k_2 and k_3 are lines of force.

always remains a potential difference between that electrode and the probing electrode caused by lower conductivity of the surface layer. It was found necessary to clean the electrodes and the partition wall daily with concentrated sulphuric acid.

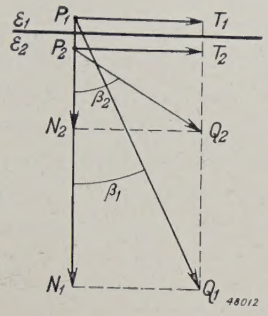


Fig. 9. P_1 and P_2 are points on either side of the boundary surface between media with different dielectric constants (ε_1, ε_2). The electrical field strength (P_1Q_1, P_2Q_2) at these points has the same tangential component ($P_1T_1 = P_2T_2$), but the normal components are inversely proportional to the dielectric constants ($P_1N_1 : P_2N_2 = \varepsilon_2 : \varepsilon_1$). From this follows the change in direction of the lines of force occurring at the boundary surface and also that of the equipotential lines intersecting the lines of force at right angles.

Fig. 10 shows some of the results of the determination of the field strength at the mercury surface with walls 0.35, 1.0 and 3.0 mm thick. Here $1/F$ (F expressed in kV/cm) is plotted as a function of the height above the boundary line between the mercury and the wall of the igniter. Since an infinitely high field strength is expected at this boundary line, due to the discontinuity caused by the limiting angle, the curve for $1/F$ is there drawn to zero.

Mechanism of the ignition

In order to explain this conception of the mechanism of the igniter we shall consider the process of ignition somewhat more closely.

It has been found by direct observation that the spark arises at a height above the mercury igniter wall which, with a given thickness and material of the wall, depends only on the voltage of the igniter. For example, with a wall thickness of 0.35 mm we found the following values for these two quantities:

3.15 kV	0.21 mm
3.55 kV	0.065 mm
4.5 kV	0.025 mm

At these voltages it follows from fig. 10, curve *a*, that the field strength at those places is: 94.5, 255 and 750 kV/cm.

The question now is whether or not cold electron emission can be expected with these field strengths. R. Haefer determined the necessary field strength with tungsten points examined ultramicroscopically for smoothness of surface [8]). In agreement with the theoretically derived values he found that for a current density of 45 A/cm^2 a field strength of 3×10^4 kV/cm is required, and that for 10^4 kV/cm only 3×10^{-16} A/cm^2 may be expected. The same applies to mercury, which has almost the same work function as that of tungsten. Consequently the field strengths of 10^2-10^3 kV/cm occurring in our case are unable to cause any appreciable cold emission.

The following two points now require an explanation:

1) How can cold emission occur at the place observed while the field strength deduced from the measurements is too low?
2) Why is it that the spark does not arise deeper in the gap between meniscus and quartz wall, where the field strength is higher?

The answer to the last question would lead us too far afield; suffice it to state that deeper in the gap some discharge can indeed be observed, but this cannot develop into a cathode spot, because the potential difference passed through by the electrons emitted is too small.

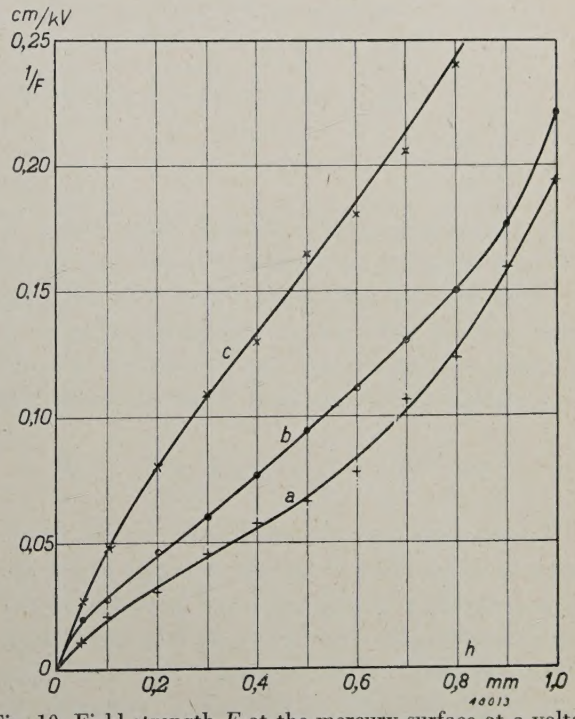


Fig. 10. Field strength F at the mercury surface at a voltage of 1 kV on the igniter. $1/F$ (F in kV/cm) is plotted as a function of the height h above the mercury-quartz boundary line on which the ignition takes place (in order better to extrapolate to small values of h, $1/F$ instead of F is plotted). The curves *a*, *b* and *c* are again valid for wall thicknesses of 0.35 mm, 1.0 mm and 3.0 mm respectively.

[8]) R. Haefer, Z. Phys. **116**, 604, 1940.

As to the first point, it has been found to be very probable, also on the basis of our further investigation of igniters, that a mechanism already considered theoretically by Tonks [9]) provides the explanation. The mercury surface, instead of being absolutely smooth, has microscopically small humps, which according to Tonks can be drawn out by the electric field to such sharp points as to produce a field strength sufficiently high for cold emission. The humps are caused by thermal agitation and in the absence of an electric field are kept relatively flat by the surface tension.

Briefly the whole process resolved itself into the following.

Upon voltage being applied to the igniter the electric field draws out the originally flat humps on the mercury surface to sharp points, where the field strength reaches values at which cold emission occurs. The emerging current of electrons flows during the first moments towards the wall of the igniter and then through ionization of the mercury vapour develops into an arc discharge.

It appears that the times necessary, according to Tonks' calculations, to draw out the humps into points, at the field strengths which follow from the above described measurements in the electrolytic tank, agree well with the ignition lags measured by us. The latter thus go to confirm Tonks' theory.

[9]) L. Tonks, Phys. Rev. **48**, 562, 1935.